

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2004年5月6日 (06.05.2004)

PCT

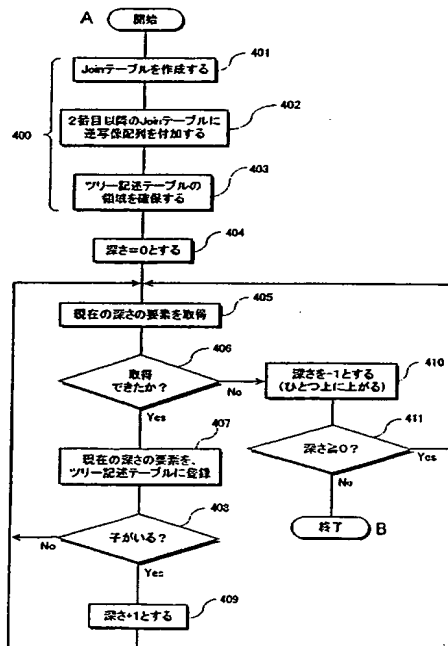
(10) 国際公開番号
WO 2004/038612 A1

- (51) 国際特許分類⁷: G06F 17/30, 12/00 (71) 出願人(米国を除く全ての指定国について): 株式会社ターボデータラボラトリー (TURBO DATA LABORATORIES INC.) [JP/JP]; 〒221-0005 神奈川県横浜市神奈川区松見町四丁目1101番地7 Kanagawa (JP).
- (21) 国際出願番号: PCT/JP2003/013558
- (22) 国際出願日: 2003年10月23日 (23.10.2003)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2002-309680
2002年10月24日 (24.10.2002) JP
- (72) 発明者; および
(75) 発明者/出願人(米国についてのみ): 古庄 晋二 (FURUSHO, Shinji) [JP/JP]; 〒221-0005 神奈川県横浜市神奈川区松見町四丁目1101番地7コートハウス菊名 804号 Kanagawa (JP).

[続葉有]

(54) Title: METHOD FOR CONVERTING CONCATENATED JOIN TABLES INTO TREE STRUCTURE AND CONVERSION PROGRAM

(54) 発明の名称: 連鎖したジョインテーブルのツリー構造への変換方法、および、変換プログラム



A...START
401...CREATE JOIN TABLE
402...ADD REVERSE MAPPING SEQUENCE TO SECOND JOIN TABLE
AND AFTER
403...RESERVE REGION OF TREE DESCRIPTION TABLE
404...SET DEPTH TO 0
405...ACQUIRE CURRENT DEPTH ELEMENT
406...ACQUIRED?
407...LOAD CURRENT DEPTH ELEMENT ON TREE DESCRIPTION
TABLE
408...CHILD PRESENT?
409...SET DEPTH TO +1
410...SET DEPTH TO -1 (RAISED BY ONE)
411...DEPTH ≥ 0
B...END

(57) Abstract: Concatenated join tables are converted into a tree structure at a high speed. Firstly, table-formatted data is divided into one or more information blocks consisting of a value list containing item values in the sequence of the item value numbers corresponding to item values belonging to particular items and a pointer sequence containing pointer values for indicating the item value numbers in the sequence of the unique record number. Next, table-formatted data whose items are to be made common are selected, value lists of a predetermined item are made equivalent, and join is executed. Join is repeated between necessary table-formatted data. Next, among the concatenated table data, one which is to be a root is selected. According to the table-formatted data concatenation, depth of the table-formatted data is decided. According to the table-formatted data concatenation and depth, a value indicating a record is arranged on a tree description table describing the tree.

(57) 要約: 連鎖したジョインテーブルをツリー構造に、高速変換する。まず、各表形式データを、各々が、特定の項目に属する項目値に対応した項目値番号の順に当該項目値が格納されている値リストと、一意的なレコード番号の順に、当該項目値番号を指示するためのポインタ値が格納されたポインタ配列とからなる一以上の情報ブロックに分割する。次いで、項目を共通化すべき表形式データを選択して、所定の項目の値リストを等価にして、ジョインを実行する。ジョインは、必要な表形式データ間で繰り返される。次いで、結合された表形式データのうち、根となるものを選択し、表形式データの結合にしたがって、各表形式データの深さを決定し、表形式データの結合および深さにしたがって、レコードを示す値を、ツリーを記述するためのツリー記述テーブルに配置していく。



(74) 代理人: 吉田 聡 (YOSHIDA, Satoshi); 〒233-0001 神奈川県 横浜市 港南区上大岡東 2-2 4-4 Kanagawa (JP).

添付公開書類:
— 国際調査報告書

(81) 指定国 (国内): CA, CN, KR, US.

(84) 指定国 (広域): ヨーロッパ特許 (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR).

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

明 細 書

連鎖したジョインテーブルのツリー構造への変換方法、および、変換プログラム

発明の属する技術分野

- 5 本発明は、連鎖したジョインテーブルをツリー構造に、高速変換する手法に関する。

背 景 技 術

- データベースは種々の用途に用いられているが、中規模ないし大規模システムにおいては、論理的な矛盾が排除できるリレーショナルデータベース（RDB）
10 の使用が主流となっている。RDBは、基本的には、項目および値と、当該値と関連付けられた他の項目の値へのポインタとから構成されているため、これをツリー構造化することは容易ではなかった。特に、大量のデータを取り扱う必要がある場合に、RDBでは大量のリレーションが発生し、データ量が著しく大きくなり、かつ、処理にも著しい時間を要する。

- 15 これに対して、XMLデータベースなどオブジェクト指向データベースが提案されている。このオブジェクト指向データベースは、データをそもそもツリー構造で保存し、その木（ツリー）の枝を順次たどっていくことが可能となる。

しかしながら、従来のデータベースにおけるツリー構造においては、以下の挙げるような問題点があった。

- 20 （１）根（ルート）が固定される。

- 従来のツリー構造では、特定のノードから他の特定のノードへのポインタで構成されているため、ノード間のリンクの繋ぎ代えが不可能である。また、ポインタを１方向にしかたどることができない。このため、ツリーは、論理的には、どの要素を根（ルート）としても良いはずであるにもかかわらず、実際には、根が
25 固定されている。

- （２）ツリー構造が固定される。

各ノードは、様々な属性を有している。たとえば、職場の住所と、住まいの住

所という項目がある場合に、職場の住所に基づくツリー（たとえば、日本一東京都一港区・・・）というツリーと、住まいの住所に基づくツリー（たとえば、日本一神奈川県一横浜市・・・）とは、論理的に別々のツリーを構成する。しかしながら、従来のデータベースにおけるツリー構造においては、これら別々のツリーを構築することが不可能であった。

（３）処理速度が著しく遅く、また、ソートや集計が困難である。

検索はツリーを順次たどる必要があるため著しい処理時間を要する。また、同様にツリーを順次たどって要素を特定する必要があるため、ソートや集計が困難である。

- 10 その一方、本発明者は、国際公開WO 00 / 1 0 1 0 3号パンフレット（第3図および第4図）に記載したような、表形式データを、各々が、特定の項目に属する項目値に対応した項目値番号の順に当該項目値が格納されている値リストと、一意的なレコード番号の順に、当該項目値番号を指示するためのポインタ値が格納されたポインタ配列とからなる一以上の情報ブロックに分割し、レコード番号、
- 15 ポインタ配列中、当該レコード番号に対応する位置の要素（ポインタ値）、および、値リスト中、ポインタ値が指し示す位置の要素（項目値）を順次たどることにより、レコード番号に対応する項目値を特定する技術を考案している。

このようなデータ形式を採用することにより、従来のRDBと比較して、検索、集計、ソートなどを著しく高速に実現することができた。

- 20 また、特開2000-339390号公報（第（14）頁～第（17）頁および図10ないし図18）に開示したように、上記データ形式を採用した複数のテーブルをジョインする手法も、本発明者は提案している。

- これにより、項目値が共通する項目をキーとして、複数のテーブルをジョインすることが可能となり、かつ、当該ジョインされたテーブルの検索、集計なども
- 25 容易に実現することができる。

本発明者は、WO 00 / 1 0 1 0 3号公報や特開2000-339390号公報に開示された技術を利用し、1以上のツリー構造を構築することができることに知見し、これを実現した。この方法およびプログラムによれば、上述した

(1) ~ (3) の問題点が解決され、任意の根（ルート）から、種々の項目のツリー構造を構築できた。かつ、その処理速度は高速であり、また、集計やソートを実現することもできた。

つまり、本発明は、任意の根を選択して、所望の 1 以上のツリー構造を高速に構築することができ、検索、集計、ソートが高速に実行できる方法およびプログラムを提供することを目的とする。

発 明 の 開 示

本発明の目的は、各々が、項目とこれに含まれる項目値とを含むレコードの配列として表わされる複数の表形式データを結合し、結合された表形式データであるジョインテーブルをツリー構造に変換する方法であって、

- a) 各表形式データを、各々が、特定の項目に属する項目値に対応した項目値番号の順に当該項目値が格納されている値リストと、一意的なレコード番号の順に、当該項目値番号を指示するためのポインタ値が格納されたポインタ配列とからなる一以上の情報ブロックに分割するように構成するステップと、
- b) 項目を共通化すべき二つの表形式データを選択するステップと、
- c) 当該選択された二つの表形式データにおいて、共通化すべき項目を見出すステップと、
- d) 前記特定された情報ブロックのうち、デフォルトのソート順が反映される表形式データをマスタ表形式データと決定し、それ以外の表形式データをスレイブ表形式データと決定するステップと、
- e) 前記選択された表形式データにおいて、前記特定された情報ブロックに含まれる値リストを相互に等価にした場合に、マスタ側のレコードに対応するスレイブ側のレコードを対応付けるステップと、
- f) 前記二つの表形式データのうちのスレイブ側の表形式データをマスタ側とするような、他の二つの表形式データを選択し、当該他の二つの表形式データについて、前記 c) 項目を見出すステップ、d) マスタ表形式データおよびスレイブ表形式データを決定するステップ、および、e) 対応付けるステップとを実行するステップと、

g) 前記 f) のステップを、ジョインが必要な表形式データに関して繰り返すステップと、

h) ジョインにより結合された表形式データのうち、根となるものを選択し、前記表形式データの結合にしたがって、各表形式データの深さを決定するステップ

5 と、

i) 前記深さおよびレコード番号の組み合わせを値とするツリー記述テーブルのための領域を確保するステップと、

j) 初期的には最も浅い深さとなるような、ある深さのレコードを特定し、前記レコードを示す値を、深さとともに、前記ツリー記述テーブルのための領域に配

10 置するステップと、

k) 前記特定されたレコードから、スレイブ側の表形式データのレコードを特定し、当該レコードに基づき、スレイブ側の表形式データがマスタ側となるような表形式データのレコードを示す値を、前記マスタ側の表形式データの深さとともに、前記ツリー記述テーブルのための領域に配置するステップと、

15 l) 前記 k) のステップを、表形式データが存在しない深さになるまで、或いは、レコードが存在しなくなるまで繰り返すステップと、

m) 前記 j) ないし l) のステップを繰り返すことにより、ツリー記述テーブルのための領域に、所定の深さおよびレコードを特定する値を配置することにより、ツリー記述テーブルを完成させるステップとを備えたことを特徴とする方法によ

20 り達成される。

本発明によれば、二つの表形式データの組を、連鎖（チェーン）して、深さの浅い表形式データのレコードから、深い位置にある表形式データのレコードを、順次特定していく。これにより、深さとレコードを示す値とからなるツリー記述テーブルを作ることができる。ツリー記述テーブルの深さにより、表形式データ
25 が特定され、レコードを示す値により、当該表形式データの実際の値を取得することが可能となる。表形式データの選択や、チェーンのための表形式データの組み合わせ、さらには、ルートとなる表形式データの選択は、自由である。したがって、きわめて柔軟性の高いツリー構造を構築することが可能となる。

また、同じ表形式データに関して、複数の異なるツリー記述テーブルを持つことができる。したがって、同じデータについて、複数のツリー構造を構築することも可能となる。

好ましい実施態様においては、前記 e) のステップにおいて、

- 5 前記値リストを等価にする際に、マスタ側の情報ブロックにおいて、その項目値の追加にともなって、前記情報ブロックのポインタ配列を指し示すための第 2 の射影配列を生成するステップと、

前記スレイブ側の情報ブロックにおいて、共通化された項目の値の存在する数を累計化した第 3 の射影配列を生成するステップと、

- 10 スレイブ側の情報ブロックにおいて、前記共通化された項目にてソートされたレコードを指し示す値の集合である第 4 の射影配列を生成するステップとを備え、
前記第 2 の射影配列、マスタ側の情報ブロックのポインタ配列、第 3 の射影配列および第 4 の射影配列を順次たどることにより、スレイブ側のレコードが特定される。

- 15 より好ましい実施態様においては、さらに、

n) 最も浅い深さの表形式データを除き、マスタ側に関する第 2 の射影配列の逆写像配列を生成するステップを備え、

前記第 4 の射影配列からの要素が指し示す逆写像配列の要素により、次の深さの表形式データにおけるマスタ側の第 2 の射影配列の要素が特定される。

- 20 さらに、逆写像を生成するステップにおいて、

前記マスタ側に関する第 2 の射影配列の要素数と等しい要素を収容する、逆写像のための配列を配置するステップと、

前記各要素として、レコードを特定し得ない他の値を配置するステップと、

- 25 前記第 4 の射影配列の位置を要素に、かつ、要素を位置とするように、前記逆写像配列に値を付与するステップとが設けられているのが望ましい。

別の好ましい実施態様においては、さらに、前記項目を共通化すべき二つの表形式データの少なくとも一方において、前記表形式データの共通化すべき項目について、検索、集計およびソートの何れかを施すステップを備えている。

上記実施態様によれば、検索、集計、ソートなどが施された表形式データを使ったツリー構造の構築が可能となる。

このように、本発明によれば、任意の根を選択し、かつ、所望のツリー構造を作成することができる。また、ツリー構造を作成するための基となるデータとして、検索、集計、ソートを施したものを利用することも可能となる。

また、本発明の目的は、各々が、項目とこれに含まれる項目値とを含むレコードの配列として表わされる複数の表形式データを結合し、結合された表形式データであるジョインテーブルをツリー構造に変換するプログラムであって、

- a) 各表形式データを、各々が、特定の項目に属する項目値に対応した項目値番号の順に当該項目値が格納されている値リストと、一意的なレコード番号の順に、当該項目値番号を指示するためのポインタ値が格納されたポインタ配列とからなる一以上の情報ブロックに分割するように構成するステップと、
- b) 項目を共通化すべき二つの表形式データを選択するステップと、
- c) 当該選択された二つの表形式データにおいて、共通化すべき項目を見出すステップと、
- d) 前記特定された情報ブロックのうち、デフォルトのソート順が反映される表形式データをマスタ表形式データと決定し、それ以外の表形式データをスレイブ表形式データと決定するステップと、
- e) 前記選択された表形式データにおいて、前記特定された情報ブロックに含まれる値リストを相互に等価にした場合に、マスタ側のレコードに対応するスレイブ側のレコードを対応付けるステップと、
- f) 前記二つの表形式データのうちのスレイブ側の表形式データをマスタ側とするような、他の二つの表形式データを選択し、当該他の二つの表形式データについて、前記 c) 項目を見出すステップ、d) マスタ表形式データおよびスレイブ表形式データを決定するステップ、および、e) 対応付けるステップとを実行するステップと、
- g) 前記 f) のステップを、ジョインが必要な表形式データに関して繰り返すステップと、

- h) ジョインにより結合された表形式データのうち、根となるものを選択し、前記表形式データの結合にしたがって、各表形式データの深さを決定するステップと、
- i) 前記深さおよびレコード番号の組み合わせを値とするツリー記述テーブルのための領域を確保するステップと、
- 5 j) 初期的には最も浅い深さとなるような、ある深さのレコードを特定し、前記レコードを示す値を、深さとともに、前記ツリー記述テーブルのための領域に配置するステップと、
- k) 前記特定されたレコードから、スレイブ側の表形式データのレコードを特定し、当該レコードに基づき、スレイブ側の表形式データがマスタ側となるような表形式データのレコードを示す値を、前記マスタ側の表形式データの深さとともに、前記ツリー記述テーブルのための領域に配置するステップと、
- 10 l) 前記k) のステップを、表形式データが存在しない深さになるまで、或いは、レコードが存在しなくなるまで繰り返すステップと、
- 15 m) 前記j) ないしl) のステップを繰り返すことにより、ツリー記述テーブルのための領域に、所定の深さおよびレコードを特定する値を配置することにより、ツリー記述テーブルを完成させるステップとを、前記コンピュータに実行させることを特徴とするプログラムによっても達成される。

図面の簡単な説明

- 20 本発明の目的および他の目的は、添付図面とともに実施例を参照することにより、さらに明らかになるであろう。ここに、

図1は、本発明の実施の形態にかかるツリー構造の構築方法を実現できるコンピュータシステムのハードウェア構成を示すブロックダイアグラムである。

- 図2は、表形式データの例、および、当該表形式データに基づく情報ブロックの例を示す図である。
- 25

図3は、表形式データの例、および、当該表形式データに基づく情報ブロックの例を示す図である。

図4は、本実施の形態にかかるツリー構造への変換処理を示すフローチャート

である。

図 5 は、組織テーブルと社員所属テーブルとのジョインを説明する図である。

図 6 は、本実施の形態にかかる社員所属テーブルと社員履歴テーブルとのジョインを説明する図である。

- 5 図 7 は、本実施の形態にかかるジョインの連鎖（チェーン）の例を示す図である。

図 8 は、本実施の形態にかかる逆写像配列の生成を説明する図である。

図 9 は、本実施の形態にかかるツリー記述テーブルの生成に利用する種々の配列を示す図である。

- 10 図 10 は、本実施の形態にかかるツリー記述テーブルの例を示す図である。

図 11 は、第 1 の実施の形態にかかるツリー記述テーブルに値を配置する処理を説明するための図である。

図 12 は、第 1 の実施の形態にかかるツリー記述テーブルに値を配置する処理を説明するための図である。

- 15 図 13 は、第 1 の実施の形態にかかるツリー記述テーブルに値を配置する処理を説明するための図である。

図 14 は、第 1 の実施の形態にかかるツリー記述テーブルに値を配置する処理を説明するための図である。

- 20 図 15 は、第 1 の実施の形態にかかるツリー記述テーブルに値を配置する処理を説明するための図である。

図 16 は、第 1 の実施の形態にかかるツリー記述テーブルに値を配置する処理を説明するための図である。

図 17 は、第 1 の実施の形態にかかるツリー記述テーブルに値を配置する処理を説明するための図である。

- 25 図 18 は、第 1 の実施の形態にかかるツリー記述テーブルに値を配置する処理を説明するための図である。

図 19 は、第 1 の実施の形態にかかるツリー記述テーブルに値を配置する処理を説明するための図である。

図 20 は、第 1 の実施の形態にかかるツリー記述テーブルにより構築されたツリー構造を説明するための図である。

図 21 は、表形式データの他の例、および、当該表形式データに基づく情報ブロックの例を示す図である。

- 5 図 22 は、表形式データの他の例、および、当該表形式データに基づく情報ブロックの例を示す図である。

図 23 は、表形式データの他の例、および、当該表形式データに基づく情報ブロックの例を示す図である。

- 10 図 24 は、第 2 の実施の形態にかかる社員所属テーブルと社員履歴テーブルとのジョインを説明する図である。

図 25 は、第 2 の実施の形態にかかる逆写像配列の生成を説明する図である。

図 26 は、第 2 の実施の形態にかかるツリー記述テーブルの生成に利用する種々の配列を示す図である。

図 27 は、第 2 の実施の形態にかかるツリー記述テーブルを示す図である

- 15 図 28 は、第 2 の実施の形態にかかるツリー記述テーブルに値を配置する処理を説明するための図である。

図 29 は、第 2 の実施の形態にかかるツリー記述テーブルに値を配置する処理を説明するための図である。

- 20 図 30 は、第 2 の実施の形態にかかるツリー記述テーブルに値を配置する処理を説明するための図である。

図 31 は、第 2 の実施の形態にかかるツリー記述テーブルに値を配置する処理を説明するための図である。

図 32 は、第 2 の実施の形態にかかるツリー記述テーブルに値を配置する処理を説明するための図である。

- 25 図 33 は、第 2 の実施の形態にかかるツリー記述テーブルにより構築されたツリー構造を説明するための図である。

図 34 は、本発明にかかるツリー構造のさらに他の例を示す図である。

発明の好ましい実施例の説明

以下、添付図面を参照して、本発明の実施の形態につき説明を加える。図1は、本発明の実施の形態にかかる連鎖的な表形式データのジョイン、および、ジョインされたテーブル（ジョインテーブル）をツリー構造に変換するコンピュータシステム5のハードウェア構成を示すブロックダイアグラムである。図1に示すように、このコンピュータシステム10は、通常のもと同様の構成であり、プログラムを実行することによりシステム全体および個々の構成部分を制御するCPU12、ワークデータなどを記憶するRAM(Random Access Memory)14、プログラム等を記憶するROM(Read Only Memory)16、ハードディスク等の固定記憶媒体18、CD-ROM19をアクセスするためのCD-ROMドライバ20、CD-ROMドライバ20や外部ネットワーク（図示せず）と接続された外部端子との間に設けられたインタフェース（I/F）22、キーボードやマウスからなる入力装置24、CRT表示装置26を備えている。CPU12、RAM14、ROM16、外部記憶媒体18、I/F22、入力装置24および表示装置26は、バス28を介して相互に接続されている。15

本実施の形態にかかる、表形式データを連鎖的に結合（ジョイン）するプログラム、結合されたテーブル（ジョインテーブル）をツリー構造に変換するプログラムなどは、CD-ROM19に收容され、CD-ROMドライバ20に読取られても良いし、ROM16に予め記憶されていても良い。また、いったんCD-ROM19から読み出したものを、外部記憶媒体18の所定の領域に記憶しておいても良い。或いは、上記プログラムは、ネットワーク（図示せず）、外部端子およびI/F22を経て外部から供給されるものであっても良い。20

また、本実施の形態においては、ツリー構造へ的高速変換を実現するために、後述するように所定のデータ形式の情報ブロックを生成する必要がある。この情報ブロック生成プログラムも同様に、CD-ROM19に收容され、ROM16に記憶され、或いは、外部記憶媒体18に記憶されても良い。或いは、これらプログラムは、ネットワーク（図示せず）を介して、外部から供給されても良いことはいうまでもない。また、本実施の形態において、情報ブロック生成プログラ25

ムにて生成されたデータ（情報ブロック）は、RAM 14に記憶され、或いは、外部記憶媒体18の所定の領域に記憶される。

本実施の形態においては、図2および図3に示すように、表形式データ（テーブル）を、所定のデータ形式により保持している。たとえば、図2（a）に示すように、テーブル#0である組織テーブルは、実際の項目値である「組織名」が、50音順（英字ではアルファベット順）でソートされて配置された値リスト（以下、場合によって「VL」と称する。）203と、各レコード番号に対応して、当該レコードが指し示す値リスト中の番号が格納された、値リストへのポインタ配列（以下、場合によって「PVL」と称する。）202とを備えた、情報ブロック201により構成される。レコード番号に対応するPVL202中の要素が示す番号の、VL中の要素（項目値）が、当該レコードに対応する項目値となる。なお、PVL202は、テーブルのレコード数だけ要素があることが理解できるであろう。

図2（b）は、「部署」という項目の値（項目値）と、「S氏名」という項目の値（項目値）とから構成される社員所属テーブルを示している。社員所属テーブルは、情報ブロック211および情報ブロック221とから構成される。「部署」の情報ブロック211には、（「組織名」のVLと、その要素が共通する）

「部署名」がソートされて格納されたVL213と、VL213中の要素（部署名）を指し示すためのPVL212とが含まれる。また、「S氏名」の情報ブロック221は、氏名が、ソートされて格納されたVL223と、VL223中の要素（氏名）を指し示すためのPVL222とが含まれる。

同様に、図3に示すように、社員履歴テーブルも、「R氏名」という項目の値から構成される情報ブロック301と、「プロジェクト」という項目の値から構成される情報ブロック311とから構成される。図2（a）、（b）に示す例と同様に、図3の例でも、情報ブロックは、それぞれ、PVL（符号302、312参照）と、VL（303、313参照）を含む。

ここで、組織テーブルと社員所属テーブルとを参照すると、組織テーブルの「組織名」と、社員所属テーブルの「部署」とを共有できることがわかる。また、

社員所属テーブルと、社員履歴テーブルとを参照すると、社員所属テーブルの「S氏名」と、社員履歴テーブルの「R氏名」とを共有できることがわかる。本実施の形態においては、項目を共有させたジョインテーブルを連鎖的に生成することにより、ツリー構造を表す。以下、ツリー構造への変換処理につき、より詳細に説明を加える。

図4は、本実施の形態にかかるツリー構造への変換処理を示すフローチャートである。図4において、ジョイン(Join)テーブルの作成(ステップ301)、2番目以降のジョイン(Join)テーブルへの逆写像配列の付加(ステップ302)、および、ツリー記述テーブルのための領域の確保(ステップ303)が、処理の準備段階を構成する。まず、これらについて詳細に説明を加える。

なお、図2および図3に示す例では、組織テーブル、社員所属テーブル、および、社員履歴テーブルの3つのテーブルが存在する。ここでは、組織テーブルと社員所属テーブルとの間で、「組織名」(「部署」)をキーとしたジョイン(Join#0)が実行され、テーブルを共有する処理、および、社員所属テーブルと社員履歴テーブルとの間で、「S氏名」(「R氏名」)をキーとしたジョイン(Join#1)が実行される。

図5は、組織テーブルと社員所属テーブルとのジョインを説明する図である。ここでは、組織テーブルの項目値の順序が維持された状態で、組織テーブル(符号501参照)と社員所属テーブル(符号502参照)とがジョインされている。

図4にしたがって、まず、組織テーブルと社員所属テーブルとのジョインテーブルが生成される。図5(a)において、組織テーブルの項目値の順序が維持されているため、これを「マスタ側」と称し、社員所属テーブルの側を「スレイブ側」と称する。ここでは、組織テーブル501と、社員所属テーブル中の「部署」のテーブル502とがジョインされ、これがジョイン(Join#0)テーブルとなる。

図5(b)に示すように、マスタ側(符号511参照)に関しては、当該マスタ側の各行(レコード)の重複数を累計化した配列(射影配列「i」とも称する)512およびマスタ側の順序付き集合(射影配列「ii」とも称する)513

とが生成される。

射影配列「i」は、元の組織テーブル（図2（a））参照の行番号が繰り返される数（累計数）が表される。たとえば、1行目（行番号「0」）の要素は、「0」であり、2行目（行番号「1」）の要素は、「3」であることから、元の組織テーブルにおける行番号「0」に位置する「東京本店」という項目値の重複数は、「 $3 - 0 = 3$ 」であることがわかる。また、射影配列「i」において、2
5 行目（行番号「1」）の要素は「3」であり、3行目（行番号「2」）の要素は「5」であることから、元の組織テーブルにおける行番号「1」に対応する「大阪支店」という項目値の重複数は、「 $5 - 3 = 2$ 」であることがわかる。なお、
10 射影配列「i」の末尾（最下行）には、ジョインされたテーブルの行の総数（ジョインした組織テーブルでは「6」）が配置される。

射影配列「ii」は、ジョイン（共通化）した組織テーブルの行番号を指定する、元の組織テーブルの行番号が示されている。つまり、この射影配列「ii」は、共通化により項目が追加された場合に、新たな組織テーブルの行番号を指定するために利用される。射影配列「i」 5 1 2 および射影配列「ii」 5 1 3 を参照することにより、ジョインされた組織テーブルに配置される値を特定することができる。たとえば、ジョインされた組織テーブル(Join#0)に関する行番号「0」ないし「2」の3行について、順序配列「ii」の行番号「0」の要素は、「0」である。したがって、元の組織テーブルの行番号「0」の要素、つまり、「東京本店」が、ジョインされた組織テーブル中に配置されることがわかる。
20

次に、スレイブ側（符号5 2 1）においては、ジョイン(Join#0)でキーとなった項目、つまり、「部署」の値ごとに、ジョインテーブルにおける各行の重複数を累計化した配列（射影配列「iii」とも称する）5 2 2、および、スレイブ側のジョインでキーとなった項目でソートされた順序付き集合（射影配列「iv」とも称する）5 2 3 とが生成される。
25

スレイブ側では、ジョイン(Join#0)でキーとなった項目「部署」でソートされているため、射影配列「iii」は符号5 2 2にて示すものとなる。射影配列「iii」により、ソート順にしたがって、「部署」という項目中、項目値「大阪

支店」の重複数が「 $2 - 0 = 2$ 」、項目値「東京本店」の重複数が「 $5 - 2 = 3$ 」、項目値「名古屋店」の重複数が「 $6 - 5 = 1$ 」となる。

また、射影配列「iv」中の値は、元の社員所属テーブルの行番号を示す。射影テーブル「iii」および射影テーブル「iv」により、ジョインされたテーブルの要素と、元の社員所属テーブルの要素とを対応させることができる。たとえば、射影配列「iii」の行番号「0」および行番号「1」の要素を参照することにより、ソートされた「部署」という項目の最初の要素である「大阪支店」に関して、2つの要素が存在し、それぞれ、元の社員所属テーブルの行番号「1」および「5」に対応することがわかる。

- 10 このようにして、組織テーブルと社員所属テーブルとの間で、「組織名」という項目および「部署」という項目をキーとしたジョイン(Join#0)が終了すると、社員所属テーブルと社員履歴テーブルとの間で、「S氏名」という項目および「R氏名」という項目をキーとしたジョイン(Join#1)が実行される。

- 15 図6は、社員所属テーブルと社員履歴テーブルとのジョインを説明する図である。先に説明したジョイン(Join#0)と同様に、このジョインにおいても、社員所属テーブルの「S氏名」の値の順序が維持された状態となる。したがって、社員所属テーブルがマスタ側となり、社員履歴テーブルがスレイブ側となる。図6(b)は、社員所属テーブル中の「S氏名」のテーブル601と、社員履歴テーブル中の「R氏名」のテーブル602とがジョインされ、これがジョイン(Join#1)テーブルとなる。

- 20 図6(b)は、このジョイン(Join#1)におけるマスタ側611に関する、各行(レコード)の重複数を累計化した配列(射影配列「i」)612、および、順序付き集合(射影配列「ii」)613を示す。また、図6(c)は、ジョイン(Join#1)におけるスレイブ側612に関する、ジョイン(Join#1)でキーとなった項目「R氏名」の値ごとに、ジョインテーブルにおける各行の重複数を累計化した配列(射影配列「iii」)622、および、ジョインでキーとなった項目でソートされた順序付き集合(射影配列「iv」)623を示す。これらの生成は、図5に示したジョイン(Join#0)のものと同様である。

なお、最初のジョイン(Join#0)のスレイブ側の順序付き集合（射影配列「iv」）5 2 3と、2番目のジョイン(Join#1)のマスタ側の順序付き集合（射影配列「ii」）6 1 3とについて、双方の配列において、同じ値は、同じレコードを指している。これは、対象とする値（要素）が同一のテーブル（社員所属テ
5 ブル）だからである。

このように、ジョインテーブルの作成処理（図4のステップ401）により、図7に示すような組織テーブル701と社員所属テーブル703との間に介在するジョインテーブル（組織：所属テーブル Join#1）702が作られ、社員所属
10 テーブル703と社員履歴テーブル705との間に介在するジョインテーブル（所属：履歴テーブル Join#1）704が作られる。さらに、社員履歴テーブルと、さらに、ジョインすべきテーブルが存在すれば、これらの間に介在する他のジョインテーブル(Join#2)が作られることは言うまでもない。

ジョインテーブルの作成が完了すると、最初のジョインテーブル以外のジョインテーブル、つまり、2番目以降のジョインテーブルについて、逆写像配列を生
15 成して、これを付加する（ステップ402）。逆写像配列は、配列中の値により、もとの配列を示す番号を指定するものである。

図8に示すように、逆写像配列を作成するためには、逆写像の元になる配列（写像配列「ii」）と同じレコード数の配列の領域800を確保し（図8（a）参照）、当該配列中の要素として、レコードを示す番号となり得ない値（たと
20 えば、「-1」）を与える（図8（b）参照）。次いで、元の配列（射影配列「ii」613）の値が指し示す番号の要素に、元の配列の番号（レコード番号）を配置する（図8（c）参照）。このようにして、図8（c）の右端にあるような、最終的な写像配列800を得ることが可能となる。

次いで、ツリー記述テーブルのための領域が確保される（ステップ403）。
25 図10は、ツリー記述テーブルの例を示す図である。ツリー記述テーブルは、複数の、深さ(Depth)とレコード番号(Rec No)との組み合わせから構成される。初期的には、それぞれの項目には、値が付与されず、以下に説明する処理により、順次、値が付与されていく。深さとは、ツリーノードの深さを表す。図7に示す

例においては、「組織テーブル」、「社員所属テーブル」、「社員履歴テーブル」の深さは、それぞれ、「0（ゼロ）」、「1」、「2」である。

本実施の形態においては、ツリー構造を記述したツリー記述テーブルを作成するために、図9に示すように、確保されたツリー記述テーブルのための領域と、

- 5 第1のジョイン(Join#0)、つまり、組織テーブルと社員所属テーブルとの間のジョインに関する第1の配列群910と、第2のジョイン(Join#1)、つまり、社員所属テーブルと社員履歴テーブルとの間のジョインに関する第2の配列群920と、これらの間を結びつける逆写像配列921とが利用される。

- 第1の配列群910は、第1のジョイン(Join#0)の順序付集合（射影配列
10 「ii」）911、ジョインキーのPVL912、累計数配列（射影配列「iii」）913および順序付集合（射影配列「iv」）914を含む。また、第2の配列群920も、同様に、第2のジョイン(Join#1)の順序付集合（射影配列「ii」）922、ジョインキーのPVL923、累計数配列（射影配列「iii」）924および順序付集合（射影配列「iv」）925を含む。

- 15 まず、深さが「0（ゼロ）」に初期化された後（ステップ404）、現在の深さの要素が取得される（ステップ405）。要素が取得された場合には（ステップ406でイエス(Yes)）、現在の深さの要素が、その深さとともにツリー記述テーブルに登録される（ステップ407）。上記例では、「組織テーブル」中の要素が取得される。図11において、順序付集合（射影配列「ii」）の先頭（行
20 番号「0」）の要素が、その深さ「0」とともに、ツリー記述テーブルの最初の要素として配置される（符号1000参照）。

- 次いで、「子」がいるか否かが判断される（ステップ408）。ここに、「子」とは、ある深さのテーブル（たとえば、深さ「0」のテーブル（組織テーブル））の順序付集合（射影配列「ii」）から配列をたどることで達する、次の
25 深さのテーブル（たとえば、深さ「1」のテーブル）中の順序付集合（射影配列「ii」）中の対応する要素（つまり、次の深さのテーブルのレコード番号）をいう。図11の例では、図12に示すように、順序付集合（射影配列「ii」）の行番号「0」の要素「0」は、深さ「0」のテーブル（組織テーブル）のPVL、

累計数配列（射影配列「iii」）および順序付集合（射影配列「iv」）、並びに、逆写像配列および深さ「1」のテーブル（社員所属テーブル）の順序付集合（射影配列「ii」）を順にたどっていくことで、要素が存在することがわかる。したがって、ステップ408でイエス(Yes)と判断される。

- 5 この場合には、深さがインクリメントされ（ステップ409）、当該深さ（上記例では「1」）の要素が取得され、深さ（上記例では「1」）とともに、取得された要素が、ツリー記述テーブルの次の要素として配置される（符号1001参照）。なお、最初のジョイン(Join#0)における射影配列「iv」の要素、および、これに引き続いているジョイン(Join#1)における射影配列「ii」の要素は、双方
- 10 とも、同じテーブルのレコードを指している。したがって、これらが同一であれば、同じレコードを指していることに留意されたい。

- 図12において、深さ「0」のテーブル（組織テーブル）の累計数配列（射影配列「iii」）において、行番号「1」の要素「2」と、次の行番号「2」の要素「5」とから、3個の要素が存在することがわかる（符号1201参照）。つまり、この深さ「0」において、上記累計数配列の行番号「1」の要素は、次に位置する順序付集合（射影配列「iv」）の行番号「2」、「3」、「4」の3つを指し示していることがわかる。なお、図12以降、累計数配列（射影配列「iii」）の右側に付与された矢印は、矢印の先に位置する要素と、矢印の元に位置する要素との差だけ、矢印の元に位置する要素が指し示すことができる、順序付集合（射影配列「iv」）の要素が存在することを意味している。
- 15
- 20

- 次いで、さらに深さがインクリメントされ（ステップ409）、次の深さの要素が取得される（ステップ405）。図12の例に関しては、深さがインクリメントされて「2」となり、深さ「1」のテーブルの順序付集合（射影配列「ii」）、PVL、累計数配列（射影配列「iii」）および順序付集合（射影配列「iv」）をたどる。図13に示すように、順序付集合（射影配列「ii」）の行番号「0」の要素をたどると、最終的に、順序付集合（射影集合「iv」）の行番号「0」、「1」のそれぞれの要素「0」、「7」が取り出される。これは、累計数配列（射影配列「iii」）の関連する要素が「0」で、かつ、次の行番号の要
- 25

素が「2」であるため、要素が「 $2 - 0 = 2$ 」だけあることがわかるからである。したがって、これら順序付集合（射影配列「iv」）の行番号「0」および「1」の要素、「0」および「7」が、それぞれ、深さ「2」とともに、ツリー記述テーブルの要素として配置される（符号1002、1003参照）。

- 5 これ以上の深さ（深さ「3」）は存在しないため、深さ方向の探索が終了する。つまり、ステップ405～ステップ409のループにおいて、次のステップ406においてノー(No)と判断され、深さがデクリメントされる（ステップ410）。つまり、ひとつ浅い深さに戻り、当該深さが「0」以上であれば処理が続行される（ステップ411参照）。

- 10 上記例では、深さが「2」から「1」になるため、当該深さ「1」の要素が再度取得される。図12の符号1201に関連して説明したように、深さ「1」の累計数配列（射影配列「iii」）に関して、3つの要素が存在することがわかっている。これらのうち、2つ目の要素および3つ目の要素については未処理である。したがって、次のステップ405においては、上記2つめの要素に基づいて
- 15 処理が実行される。図14に示すように、順序付集合（射影配列「iv」）の2つ目の要素（行番号「2」の次に位置する行番号「3」の要素）から、逆写像配列、および、次の順序付集合（射影配列「ii」）がたどられ、対応する要素が取得される。これが、深さ「1」とともに、ツリー記述テーブルの要素として配置される（符号1004参照）。

- 20 さらに、深さがインクリメントされて（ステップ409参照）、深さ「2」の要素を取得する処理が実行される。図15に示すように、順序付集合（射影配列「ii」）、PVL、累計数配列（射影配列「iii」）および順序付集合（射影配列「iv」）をたどることで、深さ「2」の要素が取得される。なお、累計数配列に配置された矢印から理解できるように、ここでは、2つの要素が取得される。
- 25 取得された要素は、それぞれ、深さ「2」とともに、ツリー記述テーブルの要素として配置される（符号1005、1006）。なお、深さ「3」は存在しないため、深さ方向の探索は終了する。

同様に、図16に示すように、深さ「1」の3つ目の要素（順序付集合（射影

配列「iv」の行番号「4」の要素）から、逆写像配列および順序付集合（射影配列「ii」）をたどることで、深さ「1」の要素を取得することができる。これも、その深さ「1」とともにツリー記述テーブルに配置される（符号1007参照）。さらに、図17に示すように、深さ「2」に関しても、配列をたどることにより、

5 要素が取得され、その深さ「2」とともに、ツリー記述テーブルに配置される（符号1008参照）。なお、ここでも、深さ「3」は存在しないため、深さ方向の探索は終了する。

このような終了の後、深さをデクリメントして「1」としても要素が取得できないため（ステップ410、411、415および406）、さらに深さがデ

10 クリメントされて「0」とされて、深さ「0」の要素が取得される。図18に示すように、順序付集合（射影配列「ii」）の次（つまり、行番号「1」）の要素「1」が、ツリー記述テーブルに登録すべき要素となる。したがって、深さ「0」とともに、これがツリー記述テーブルに配置される（符号1009参照）。その後、図4、および、図11ないし図18を参照して説明したものと同様に、

15 深さ方向への順次探索、および、要素のツリー記述テーブルへの配置が繰り返される。

図19は、このような処理を続けた結果、最終的なツリー記述テーブルの要素が取得された状態を示す図である。ここに示す処理は、図示しない先の処理にて、深さ「1」の要素をツリー記述テーブルに配置したことに引き続いて実行される。

20 ツリー記述テーブルに配置された要素は、当該深さ「1」に関する「Join#1」側の順序付集合（射影配列「ii」）と共通の意味を持つ。したがって、この要素をたどることで、深さ「2」の要素を探索することができる。

図19の例では、順序付集合（射影配列「ii」）の行番号「2」の要素「2」から、PVLおよび累計数配列（射影配列「iii」）をたどり、順序付集合（射影配列「iv」）の行番号「4」の要素である「4」が取得され、これが、深さ

25 「2」とともに、ツリー記述テーブルに配置される。

このようにして完成したツリー記述テーブル900の意味につき、図20を参照して説明する。前述したように、ツリー記述テーブル900には、深さ

(Depth) とレコード番号 (Rec No) との組み合わせが配置されている。たとえば、深さの項目の値は、それがどの深さのテーブルに関連するかを特定し、かつ、レコード番号は、深さにて特定されたテーブルのレコード番号を特定している。

- たとえば、ツリー記述テーブルの最初の要素の組み合わせ (0, 0) により、
- 5 深さ「0」のテーブル (組織テーブル) の、レコード番号「0」の要素、つまり、「東京本店」が示される。また、これに引き続く要素の組み合わせ (1, 0) により、深さ「1」のテーブル (社員所属テーブル) のレコード番号「0」の要素、つまり、「東京本店」および「A氏」を示している。同様に、さらに引き続く要素の組み合わせ (2, 0)、(2, 7) は、それぞれ、深さ「2」のテーブル
- 10 (社員履歴テーブル) の、それぞれ、レコード番号「0」および「7」の要素、つまり、「A氏」および「Prj-1」、並びに、「A氏」および「Prj-4」を示すことになる。このように、ツリー記述テーブルは、複数のジョインされたテーブルをツリー構造にて記述したことが理解できるであろう。

- このように、本実施の形態によれば、所望のテーブルを指定することで、上述
- 15 した処理により、ツリー記述テーブルを作成し、ジョインされたテーブルをツリー構造にて記述することが可能となる。

- 次に、本発明の第2の実施の形態につき説明を加える。第1の実施の形態においては、表形式データの全体集合をジョインしたもののツリー記述テーブルを生成したが、第2の実施の形態では、一部の表形式データを、所定の項目で検索し、
- 20 かつ、ソートした部分集合を用いている。なお、第2の実施の形態においても、システムのハード構成 (図1) およびツリー構造への変換処理の概略 (図4) については、第1の実施の形態のものとほぼ同様である。

- 図21～図22は、第2の実施の形態にかかる表形式データ (テーブル) の例を示す図である。第2の実施の形態においては、組織テーブルおよび社員履歴テーブルは、第1の実施の形態と同様のものを利用し (図21の符号201および図23の符号301, 311参照)、かつ、ジョインの対象は全体集合としている (図21の符号2101および図23の符号2301参照)。その一方、社員所属テーブル自体は、第1の実施の形態のものと同様であるが (図22の符号2
- 25

21

11、221参照)が、テーブルを「部署」の項目の「大阪支店」、「名古屋支店」でOR検索し(図22の符号2202参照)、かつ、項目「S氏名」でソートしている(図22の符号2203参照)。

- 5 これら表形式データの間で、「組織名」(「部署」)をキーとしたジョイン(Join#0)は、全体集合(符号2101、2201参照)の間のジョインであり、その一方、「S氏名」(「R氏名」)をキーとしたジョイン(Join#1)は、検索およびソートされた部分集合(図22の符号2203参照)と全体集合(図23の符号2301参照)との間のジョインとなる。

- 10 「組織名」(「部署」)をキーとしたジョイン(Join#0)に関して作成される配列は、全体集合に関するものであるため、図5に示すものと同様になる。その一方、「S氏名」(「R氏名」)をキーとしたジョイン(Join#1)は、図24に示すようなものとなる。図24に示すように、このジョイン(Join#1)では、項目「部署」の「大阪支店」、「名古屋店」でOR検索され、項目「S氏名」でソートされた社員所属テーブル(マスタ側)2401、および、上記OR検索により抽出された「S氏名」と同一の「R氏名」をもつような社員履歴テーブル(スレ
- 15 イブ側)2402に関連して種々の配列が作成される。

- 図24(b)に示すように、マスタ側に関しては、マスタ側の各行の重複数を類型化した射影配列「i」2412、および、マスタ側の順序付き集合である射影配列「ii」2413が生成される。その一方、図24(c)に示すように、ス
- 20 レイブ側に関しては、ジョイン(Join#1)でキーとなった項目、つまり、「R氏名」の値ごとに、ジョインテーブルにおける各行の重複数を累計化した配列である射影配列「iii」2422、および、上記キーとなった項目でソートされた順序付き集合である射影配列「iv」2423が生成される。これら各種配列およびその要素が示す意味については、図5および図6を参照して説明した、第1の実
- 25 施の形態のものと同様である。

次いで、逆写像配列の作成(図4のステップ402)につき説明を加える。逆写像配列の作成も、基本的には第1の実施の形態と同様である。ただし、逆写像配列の要素数は、社員所属テーブル(マスタ側)の全レコード数に対応する(図

25 (a) 参照)。次いで、図25 (b) に示すように、第1の実施の形態と同様に、各要素に、レコード番号となり得ない値（たとえば、「-1」）を付与し、その後、図25 (c) に示すように、元の配列の値が指し示す番号の要素に、元の配列の番号（レコード番号）を配置する。これにより、図25 (c) の右端に示すような逆写像配列2500を得ることができる。第2の実施の形態においては、逆写像配列の要素中、レコード番号を示していないものが含まれることに留意すべきである。

このようにして、種々の射影配列および逆写像配列が生成された後に、図4のステップ404～411の処理が繰り返され、ツリー記述テーブルが出来上がる。

10 図26は、ツリー記述テーブルを作成するために利用される配列群などを示す図である。図26に示すように、ツリー記述テーブルを作成するために、確保されたツリー記述テーブルのための領域と、第1のジョイン(Join#0)、つまり、組織テーブルと社員所属テーブルとの間のジョインに関する第1の配列群2610と、第2のジョイン(Join#1)、つまり、社員所属テーブルと社員履歴テーブルとの間のジョインに関する第2の配列群2620と、これらの間を結びつける逆写像配列2621とが利用される。

第1の配列群2610は、図9に示す第1の配列群910と同様であり、第1のジョイン(Join#0)の順序付集合（射影配列「ii」）2611、ジョインキーのPVL2612、累計数配列（射影配列「iii」）2613および順序付集合（射影配列「iv」）2614を含む。また、第2の配列群2620も、同様に、第2のジョイン(Join#1)の順序付集合（射影配列「ii」）2622、ジョインキーのPVL2623、累計数配列（射影配列「iii」）2624および順序付集合（射影配列「iv」）2625を含む。

図27に示すように、ツリー記述テーブル2600においては、深さ(Depth)およびレコード番号(Rec No)の組み合わせが、順次配置されていく。

ツリー記述テーブルに要素を配置する処理においては、まず、深さが「0（ゼロ）」に初期化された後（図4のステップ404）、現在の深さの要素が取得され、これが、深さとともに配置される（ステップ405～407）。図28に示

すように、初期的には、第1のジョイン(Join#0)の順序付集合(射影配列「ii」)の行番号「0」の要素が、ツリー記述テーブル2600の最初の要素として配置される(符号2700参照)。

次いで、図4にしたがって、「子」がいるか否かが判断される(ステップ408)。図29に示すように、第1のジョイン(Join#0)に関して、順序付集合(射影配列「ii」)の行番号「0」の要素から、PVL、累計数配列(射影配列「iii」)および順序付集合(射影配列「iv」)の対応する要素をたどった結果、指し示される逆写像配列の3つの要素がすべて「-1」となる。前述したように逆写像配列の要素「-1」は、レコードが存在しないことを示しているため、この場合には、第2のジョイン(Join#1)に関して、「子」となる要素が存在していないことになる(符号2901参照)。

したがって、図4の処理においては、ステップ408でノー(No)と判断され、同じ深さ「0」の次の要素が取得される(ステップ405参照)。上記例では、図30に示すように、順序付集合(射影配列「ii」)の行番号「1」の要素「1」が、その深さ「0」とともにツリー記述テーブルに配置される(符号2701参照)。

次いで、「子」となる要素の有無が判断される。上記順序付集合(射影配列「ii」)の要素から、PVL、累計数配列(射影配列「iii」)、順序付集合(射影配列「iv」)、および、逆写像配列の対応する要素をたどった結果、第2のジョイン(Join#1)の順序付集合(射影配列「ii」)の要素が特定される。したがって、この順序付集合(射影配列「ii」)の行番号「1」の要素「1」が、その深さ「1」とともに、ツリー記述テーブルに配置される(符号2702参照)。

さらに、「子」となる要素の有無が判断される。上記例では、第2のジョイン(Join#1)の順序付集合の要素から、PVL、累計数配列(射影配列「iii」)、および、順序付集合(射影配列「iv」)の要素が特定され、順序付集合(射影配列「iv」)の2つの要素が、それぞれ、その深さ「2」とともに、ツリー記述テーブルに配置される(符号2703、2704参照)。

このような処理を繰り返すことにより、図27に示すようなツリー記述テーブ

ルを完成させることができる。このようにして完成されたツリー記述テーブルの意味も、図20を参照して説明した第1の実施の形態のものと同様である。図33に示すように、ツリー記述テーブルの深さの項目の値は、どの深さのテーブルに関連するかを特定し、かつ、対応するレコード番号は、深さにて特定されたテーブルのレコード番号を特定している。

図33において、たとえば、最初の要素の組み合わせ(0, 0)により、深さ「0」のテーブル(組織テーブル)の、レコード番号「0」の要素、つまり、「東京本店」が示される。また、これに引き続く、要素の組み合わせ(0, 1)により、深さ「0」のテーブル(組織テーブル)の、レコード番号「1」の要素、つまり、「大阪支店」が示される。同様に、さらに引き続く要素の組み合わせ(1, 1)および(2, 1)により、深さ「1」のテーブル(社員所属テーブル)および深さ「2」のテーブル(社員履歴テーブル)の、それぞれ、レコード番号「1」および「1」の要素、つまり、「大阪支店」および「D氏」、並びに、「D氏」および「Prj-1」が示される。

これら実施の形態によれば、任意のノードをルートに設定して、ツリー構造を構築することが可能となる。たとえば、第2の実施の形態においては、図34の符号3400に示すようなツリー構造を構築することができた。本実施の形態を利用することで、任意の項目をルートとしてツリー構造を再構築することができる。たとえば、図34のうち、社員所属テーブルの行番号「1」のレコードである「大阪支店」および「D氏」をルートと選択して、新たなツリー構造3410を構築することも可能である。

ツリー構造を構築する場合には、親ノード群となるレコード群(表形式データ)と、子ノード群となるレコード群(表形式データ)とを選択し、これらをジョインさせればよい。これを所望の回数だけ連鎖的に繰り返した後、ツリー記述テーブルを作成することにより、任意の深さのツリー構造を構築することが可能となる。

第1および第2の実施の形態によれば、非常に高速に、複数の表形式データをジョインさせることで、ツリー構造を構築することができる。たとえば、「Pe

anium (登録商標) 4」プロセッサ (1.8 GHz) を利用すると、ジョインに関する種々の配列の作成を除き、1000万行/秒ほどの速度で、ツリー記述テーブルに、値を配置することが可能となる。これは、従来の100~1000倍程度の速さである。

- 5 また、図34を参照して説明したように、任意のノードをルートとしたツリー構造を構築することができ、非常に高い柔軟性を備えている。また、ツリー記述テーブルによりツリー構造が表されるため、同じ表形式データによる複数の異なるツリー構造を保持することが可能となる。たとえば、社員名簿を、所属部署によるツリーで表現するとともに、出生地域によるツリーで表現することが可能となる。
- 10 なる。

さらに、ソートや集計を組み合わせることができる (第2の実施の形態参照)。

- また、従来のツリー構造と比較して、使用するメモリ領域をコンパクトにすることが可能となる。たとえば、XMLデータベースにおいては、ノードの接続情報にパス名を表す文字列を埋め込むため、著しく多量なメモリ領域を必要とする。
- 15 その一方、本実施の形態によれば、ツリー記述テーブルによりツリー構造を記述することができる。本発明者の実験によれば、株価情報に関して、XMLデータベースと比較して、1/5のメモリ量にてツリー構造を表現できた。

- 本発明は、以上の実施の形態に限定されることなく、特許請求の範囲に記載された発明の範囲内で、種々の変更が可能であり、それらも本発明の範囲内に包含
- 20 されるものであることは言うまでもない。

たとえば、前記実施の形態においては、3つの表形式データを、連鎖的にジョインしているが、連鎖の数はこれに限定されるものでないことは言うまでもない。

- また、第1の実施の形態において、あるジョインと引き続くジョインとの間に、逆写像配列を介在させている。しかしながら、表形式データの全体集合をそのまま利用する場合には、この逆写像配列を省略しても良い。
- 25 ま利用する場合には、この逆写像配列を省略しても良い。

さらに、前記実施の形態においては、一般のコンピュータシステム10内に、所定のプログラムを読み込み、当該プログラムを実行することにより、複数の表形式データのジョインおよびジョインされた表形式データに関する処理を実現し

ているが、本発明はこれに限定されるものではなく、パーソナルコンピュータ等のような一般のコンピュータシステムに、データベース処理専用のボードコンピュータを接続し、当該ボードコンピュータが上記処理を実行できるように構成しても良いことは言うまでもない。したがって、本明細書において、手段とは必ずしも物理的手段を意味するものではなく、各手段の機能が、ソフトウェアおよび/またはハードウェアによって実現される場合を包含する。さらに、一つの手段の機能が、二つ以上の物理的手段により実現されても、若しくは、二つ以上の手段の機能が、一つの物理的手段により実現されてもよい。

本発明は、たとえば、RDBにて保持される部品リストに基づき、部品表（BOM）を作成するための処理（BOM展開）などに利用することができる。また、本発明は、分流や合流を特徴とするフローにおいて、これを体系的に手際よく管理することが可能となる。たとえば、

ある食品（たとえば、畜肉）が、一頭ごとにどの部位がどのような経路を経て、そこで販売されたかを管理する場合に、これはツリー状のデータとなる。一般に、第1次産業や第2次産業において作られた生産物は、消費されるか、或いは、形を変えるまで、ツリー状に拡散をしていく。本発明は、このように、分流する状態を示すツリー構造を構築し、これを管理することが可能となる。また、本発明によれば、時間の経過に伴いツリーが伸張する状態を、ツリー記述テーブルを作成することにより把握することが可能となる。

その一方、ある工業製品を生産するために、様々な部品、エネルギーおよびサービスを購入する場合にも、ツリー状のデータとなる。たとえば、自動車（最終製品）を生産するために、ある部品会社から点火プラグを購入し、部品会社は、電力会社から電力を購入するとともに、点火プラグの部品であるセラミックを、他の部品会社から購入する。このように、時間を遡るのにしたがって、ツリーが拡散する。本発明では、このような合流する状態を示すツリー構造を構築し、これを管理することも可能となる。

本発明によれば、任意の根を選択して、所望の1以上のツリー構造を高速に構築することができ、検索、集計、ソートが高速に実行できる方法およびプログラ

ムを提供することが可能となる。

産業上の利用分野

本発明は、特に、大量のデータを管理するシステム、たとえば、データベース、データウェアハウスに利用することできる。より具体的には、SCM
5 (Supply Chain Management)におけるBOM展開のような基幹業務管理、事務管理に好適である。

請 求 の 範 囲

1. 各々が、項目とこれに含まれる項目値とを含むレコードの配列として表わされる複数の表形式データを結合し、結合された表形式データであるジョインテーブルをツリー構造に変換する方法であって、
- 5 a) 各表形式データを、各々が、特定の項目に属する項目値に対応した項目値番号の順に当該項目値が格納されている値リストと、一意的なレコード番号の順に、当該項目値番号を指示するためのポインタ値が格納されたポインタ配列とからなる一以上の情報ブロックに分割するように構成するステップと、
- b) 項目を共通化すべき二つの表形式データを選択するステップと、
- 10 c) 当該選択された二つの表形式データにおいて、共通化すべき項目を見出すステップと、
- d) 前記特定された情報ブロックのうち、デフォルトのソート順が反映される表形式データをマスタ表形式データと決定し、それ以外の表形式データをスレイブ表形式データと決定するステップと、
- 15 e) 前記選択された表形式データにおいて、前記特定された情報ブロックに含まれる値リストを相互に等価にした場合に、マスタ側のレコードに対応するスレイブ側のレコードを対応付けるステップと、
- f) 前記二つの表形式データのうち、スレイブ側の表形式データをマスタ側とするような、他の二つの表形式データを選択し、当該他の二つの表形式データについて、前記 c) 項目を見出すステップ、d) マスタ表形式データおよびスレイブ表形式データを決定するステップ、および、e) 対応付けるステップとを
- 20 実行するステップと、
- g) 前記 f) のステップを、ジョインが必要な表形式データに関して繰り返すステップと、
- 25 h) ジョインにより結合された表形式データのうち、根となるものを選択し、前記表形式データの結合にしたがって、各表形式データの深さを決定するステップと、
- i) 前記深さおよびレコード番号の組み合わせを値とするツリー記述テーブルの

ための領域を確保するステップと、

j) 初期的には最も浅い深さとなるような、ある深さのレコードを特定し、前記レコードを示す値を、深さとともに、前記ツリー記述テーブルのための領域に配置するステップと、

5 k) 前記特定されたレコードから、スレイブ側の表形式データのレコードを特定し、当該レコードに基づき、スレイブ側の表形式データがマスタ側となるような表形式データのレコードを示す値を、前記マスタ側の表形式データの深さとともに、前記ツリー記述テーブルのための領域に配置するステップと、

1) 前記k)のステップを、表形式データが存在しない深さになるまで、或いは、
10 レコードが存在しなくなるまで繰り返すステップと、

m) 前記j)ないし1)のステップを繰り返すことにより、ツリー記述テーブルのための領域に、所定の深さおよびレコードを特定する値を配置することにより、ツリー記述テーブルを完成させるステップとを備えたことを特徴とする方法。

15 2. 前記e)のステップにおいて、前記値リストを等価にする際に、マスタ側の情報ブロックにおいて、その項目値の追加にともなって、前記情報ブロックのポインタ配列を指し示すための第2の射影配列を生成するステップと、

前記スレイブ側の情報ブロックにおいて、共通化された項目の値の存在する数を累計化した第3の射影配列を生成するステップと、

20 スレイブ側の情報ブロックにおいて、前記共通化された項目にてソートされたレコードを指し示す値の集合である第4の射影配列を生成するステップとを備え、

前記第2の射影配列、マスタ側の情報ブロックのポインタ配列、第3の射影配列および第4の射影配列を順次たどることにより、スレイブ側のレコードが
25 特定されることを特徴とする請求の範囲第1項に記載の方法。

3. さらに、

n) 最も浅い深さの表形式データを除き、マスタ側に関する第2の射影配列の逆写像配列を生成するステップを備え、前記第4の射影配列からの要素が指し示

す逆写像配列の要素により、次の深さの表形式データにおけるマスタ側の第2の射影配列の要素が特定されることを特徴とする請求の範囲第2項に記載の方法。

4. 前記逆写像を生成するステップにおいて、前記マスタ側に関する第2の射影
5 配列の要素数と等しい要素を収容する、逆写像のための配列を配置するステップと、
前記各要素として、レコードを特定し得ない他の値を配置するステップと、
前記第4の射影配列の位置を要素に、かつ、要素を位置とするように、前記
逆写像配列に値を付与するステップとを備えたことを特徴とする請求の範囲第
10 3項に記載の方法。
5. さらに、前記項目を共通化すべき二つの表形式データの少なくとも一方において、前記表形式データの共通化すべき項目について、検索、集計およびソートの何れかを施すステップを備えたことを特徴とする請求の範囲第1項ないし第3項の何れか一項に記載の方法。
- 15 6. 各々が、項目とこれに含まれる項目値とを含むレコードの配列として表わされる複数の表形式データを結合し、結合された表形式データであるジョインテーブルをツリー構造に変換するプログラムであって、
 - a) 各表形式データを、各々が、特定の項目に属する項目値に対応した項目値番号の順に当該項目値が格納されている値リストと、一意的なレコード番号の順
20 に、当該項目値番号を指示するためのポインタ値が格納されたポインタ配列とからなる一以上の情報ブロックに分割するように構成するステップと、
b) 項目を共通化すべき二つの表形式データを選択するステップと、
c) 当該選択された二つの表形式データにおいて、共通化すべき項目を見出すステップと、
25 d) 前記特定された情報ブロックのうち、デフォルトのソート順が反映される表形式データをマスタ表形式データと決定し、それ以外の表形式データをスレイブ表形式データと決定するステップと、
e) 前記選択された表形式データにおいて、前記特定された情報ブロックに含ま

れる値リストを相互に等価にした場合に、マスタ側のレコードに対応するスレイブ側のレコードを対応付けるステップと、

- f) 前記二つの表形式データのうち、スレイブ側の表形式データをマスタ側とするような、他の二つの表形式データを選択し、当該他の二つの表形式データについて、前記 c) 項目を見出すステップ、 d) マスタ表形式データおよびスレイブ表形式データを決定するステップ、および、 e) 対応付けるステップとを実行するステップと、
- g) 前記 f) のステップを、ジョインが必要な表形式データに関して繰り返すステップと、
- 10 h) ジョインにより結合された表形式データのうち、根となるものを選択し、前記表形式データの結合にしたがって、各表形式データの深さを決定するステップと、
- i) 前記深さおよびレコード番号の組み合わせを値とするツリー記述テーブルのための領域を確保するステップと、
- 15 j) 初期的には最も浅い深さとなるような、ある深さのレコードを特定し、前記レコードを示す値を、深さとともに、前記ツリー記述テーブルのための領域に配置するステップと、
- k) 前記特定されたレコードから、スレイブ側の表形式データのレコードを特定し、当該レコードに基づき、スレイブ側の表形式データがマスタ側となるような表形式データのレコードを示す値を、前記マスタ側の表形式データの深さとともに、前記ツリー記述テーブルのための領域に配置するステップと、
- 20 l) 前記 k) のステップを、表形式データが存在しない深さになるまで、或いは、レコードが存在しなくなるまで繰り返すステップと、
- m) 前記 j) ないし l) のステップを繰り返すことにより、ツリー記述テーブルのための領域に、所定の深さおよびレコードを特定する値を配置することにより、ツリー記述テーブルを完成させるステップとを、前記コンピュータに実行させることを特徴とするプログラム。

7. 前記 e) のステップにおいて、前記値リストを等価にする際に、マスタ側の

情報ブロックにおいて、その項目値の追加にともなって、前記情報ブロックのポインタ配列を指し示すための第2の射影配列を生成するステップと、

前記スレイブ側の情報ブロックにおいて、共通化された項目の値の存在する数を累計化した第3の射影配列を生成するステップと、

- 5 スレイブ側の情報ブロックにおいて、前記共通化された項目にてソートされたレコードを指し示す値の集合である第4の射影配列を生成するステップとを、前記コンピュータに実行させ、

- 前記第2の射影配列、マスタ側の情報ブロックのポインタ配列、第3の射影配列および第4の射影配列を順次たどることにより、スレイブ側のレコードが
10 特定されるよう、前記コンピュータを動作させることを特徴とする請求の範囲第6項に記載のプログラム。

8. さらに、

- n) 最も浅い深さの表形式データを除き、マスタ側に関する第2の射影配列の逆写像配列を生成するステップを、前記コンピュータに実行させ、前記第4の射影配列からの要素が指し示す逆写像配列の要素により、次の深さの表形式データにおけるマスタ側の第2の射影配列の要素が特定されるように、前記コンピュータを動作させることを特徴とする請求の範囲第7項に記載のプログラム。

9. 前記逆写像を生成するステップにおいて、前記マスタ側に関する第2の射影配列の要素数と等しい要素を収容する、逆写像のための配列を配置するステップと、
20

 前記各要素として、レコードを特定し得ない他の値を配置するステップと、前記第4の射影配列の位置を要素に、かつ、要素を位置とするように、前記逆写像配列に値を付与するステップとを、前記コンピュータに実行させることを特徴とする請求の範囲第8項に記載のプログラム。

- 25 10. さらに、前記項目を共通化すべき二つの表形式データの少なくとも一方において、前記表形式データの共通化すべき項目について、検索、集計およびソートの何れかを施すステップを、前記コンピュータに実行させることを特徴とする請求の範囲第6項ないし第9項の何れか一項に記載のプログラム。

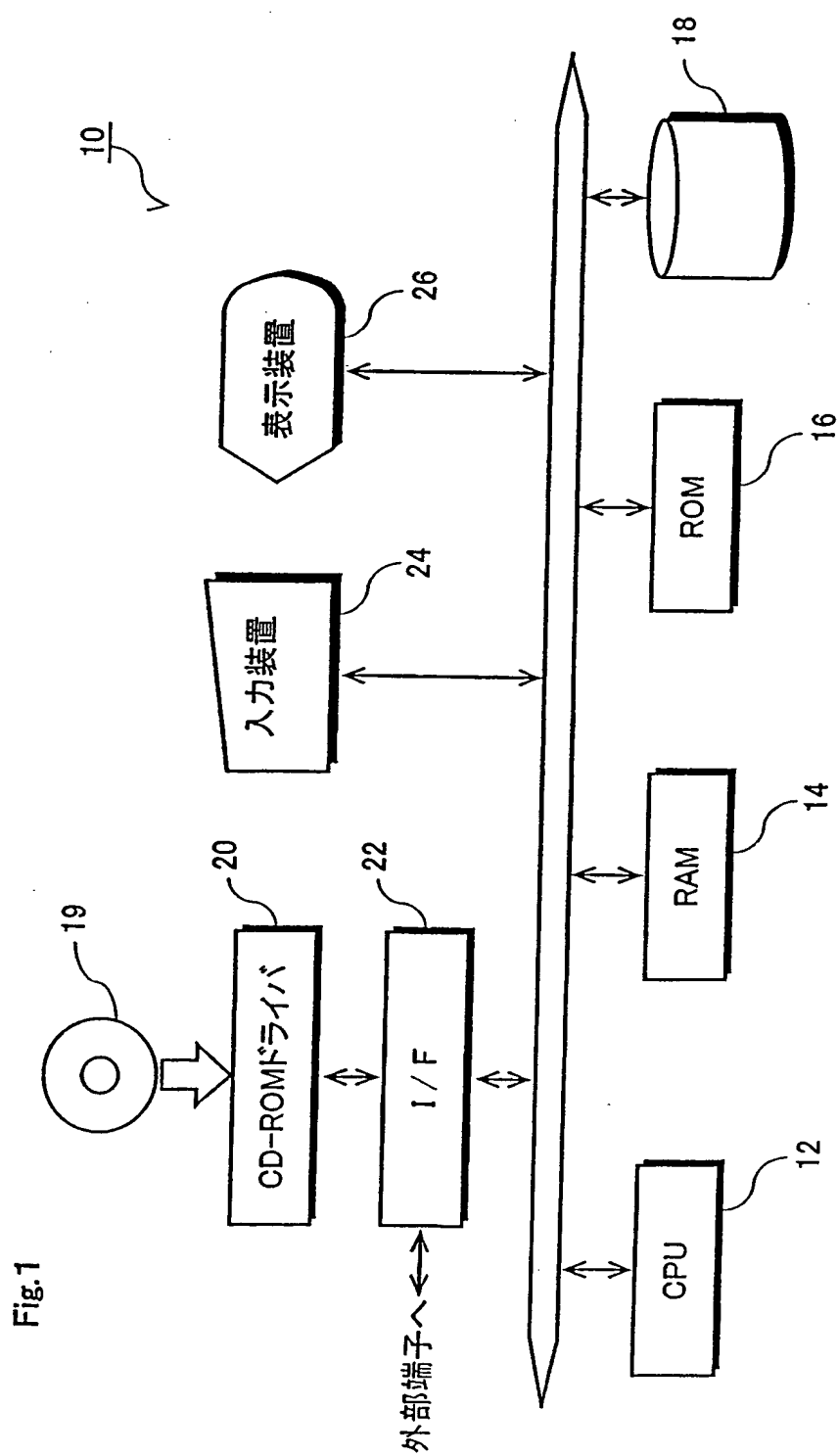


Fig.2

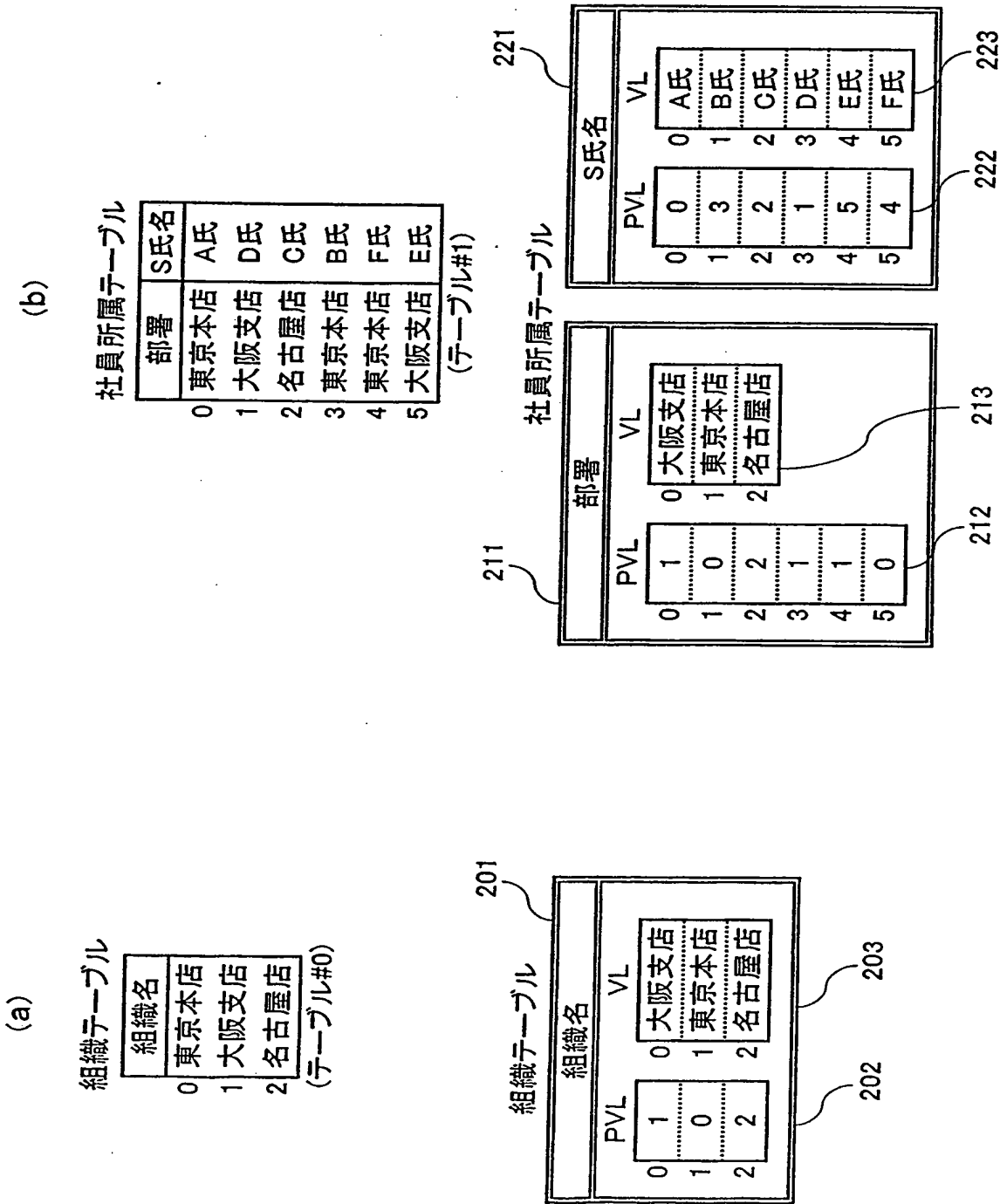


Fig.3

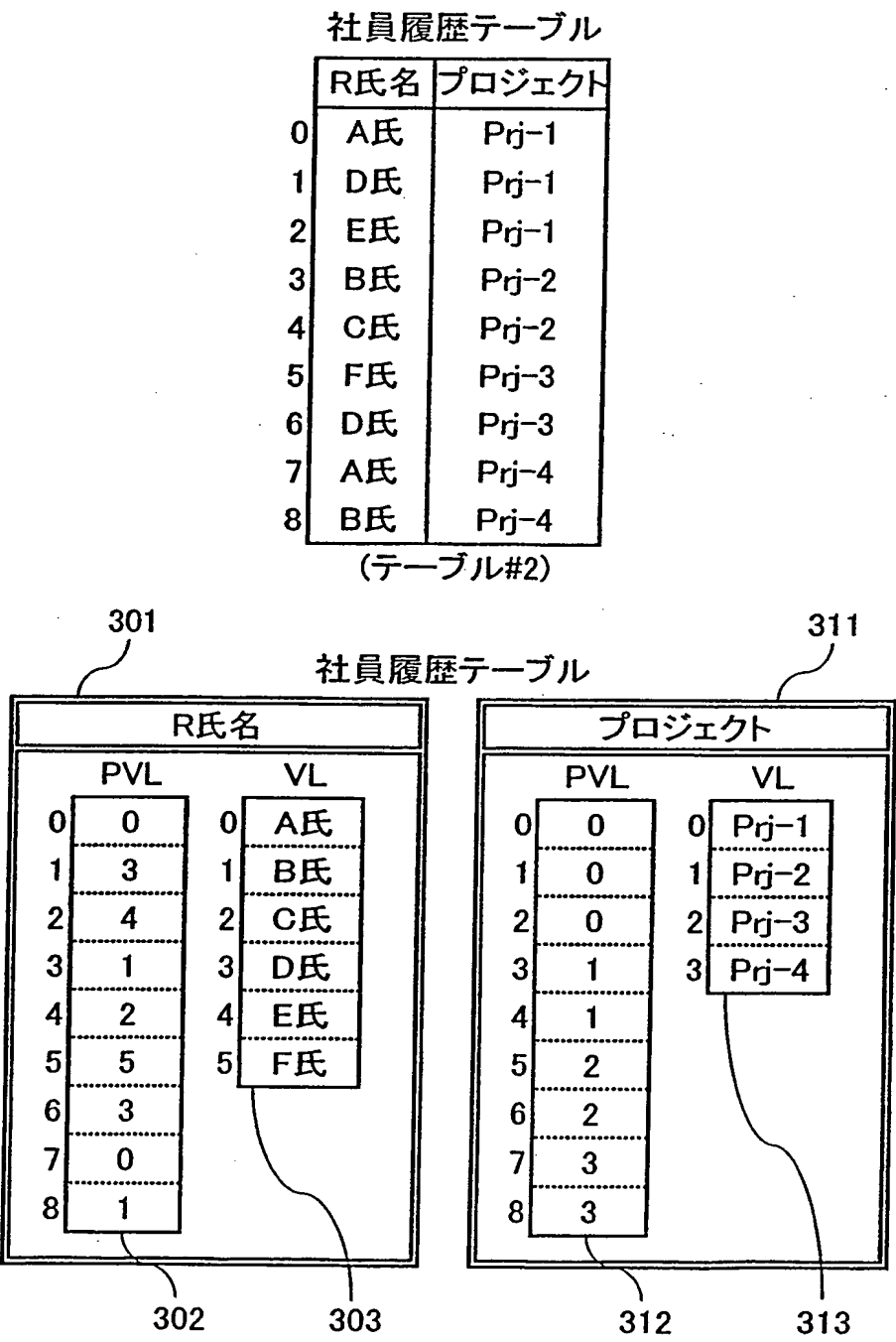


Fig.4

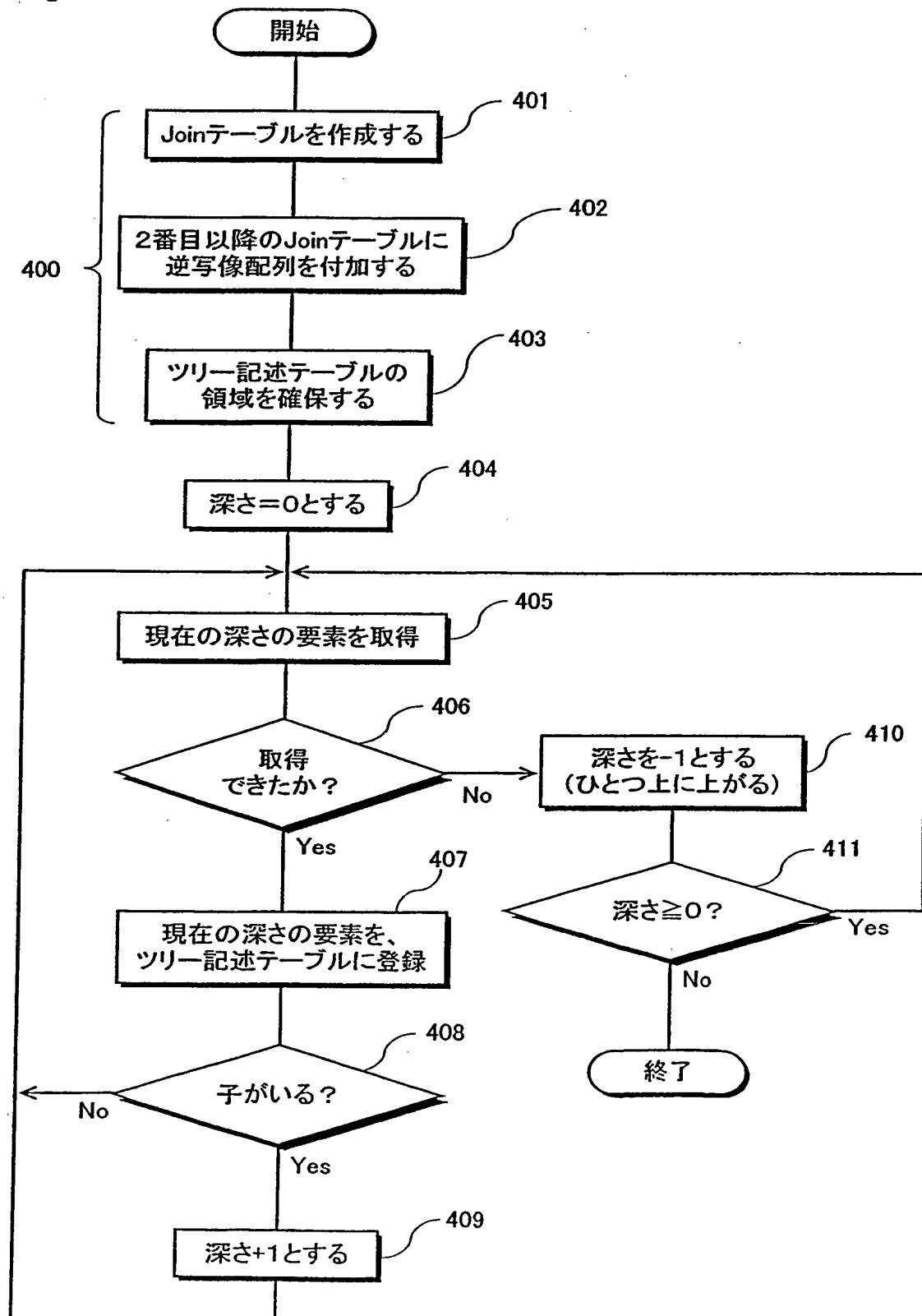
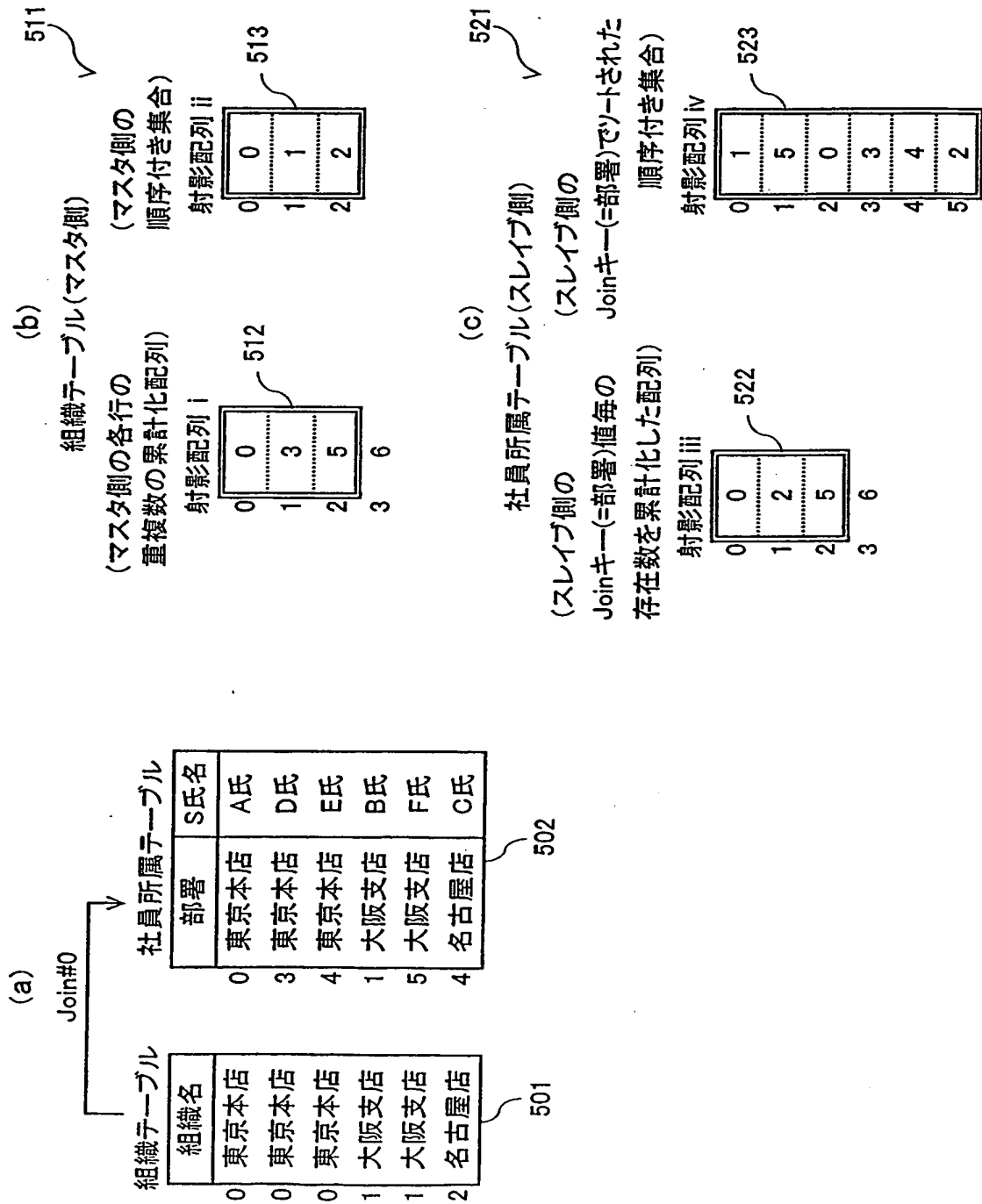
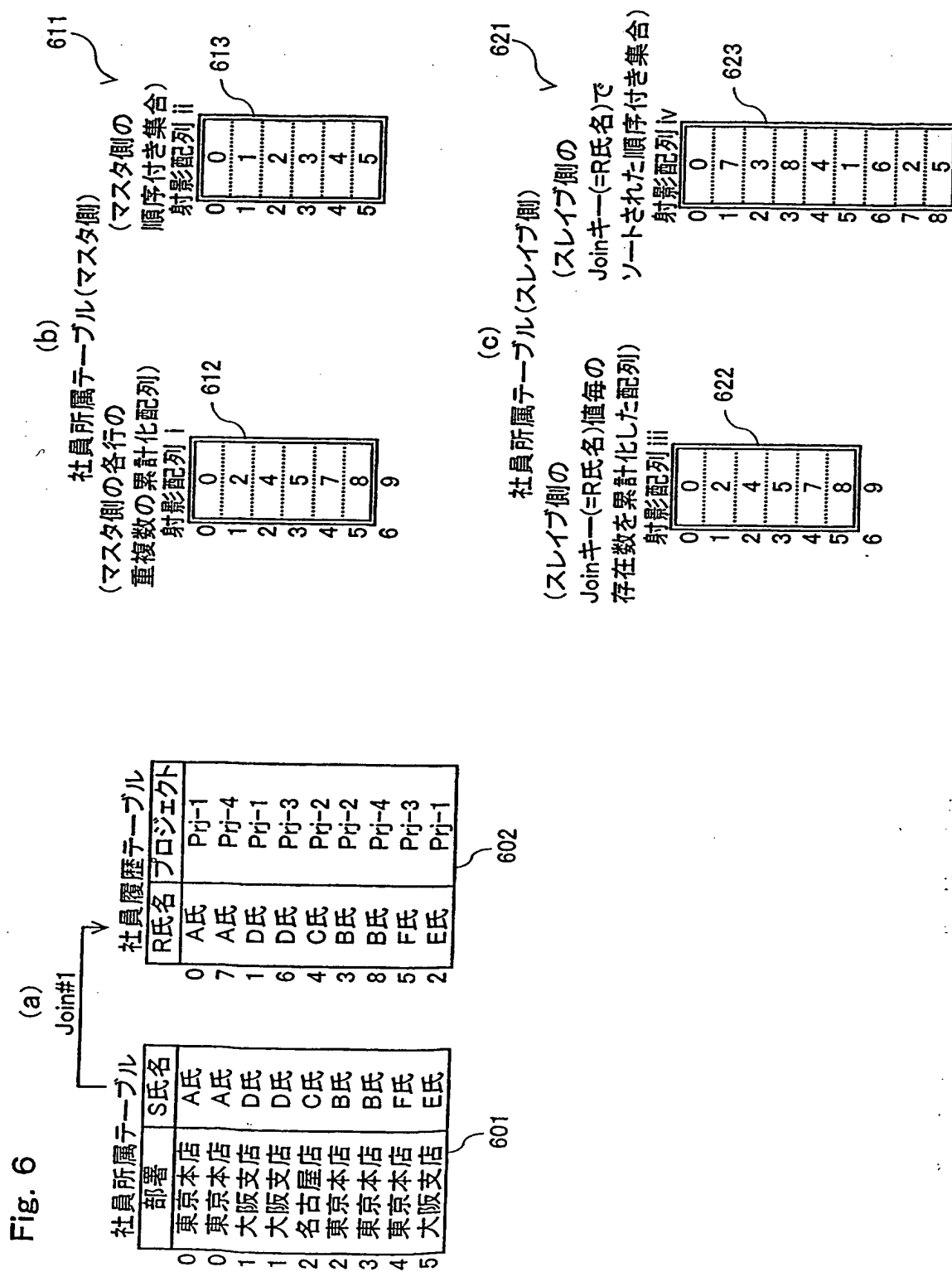
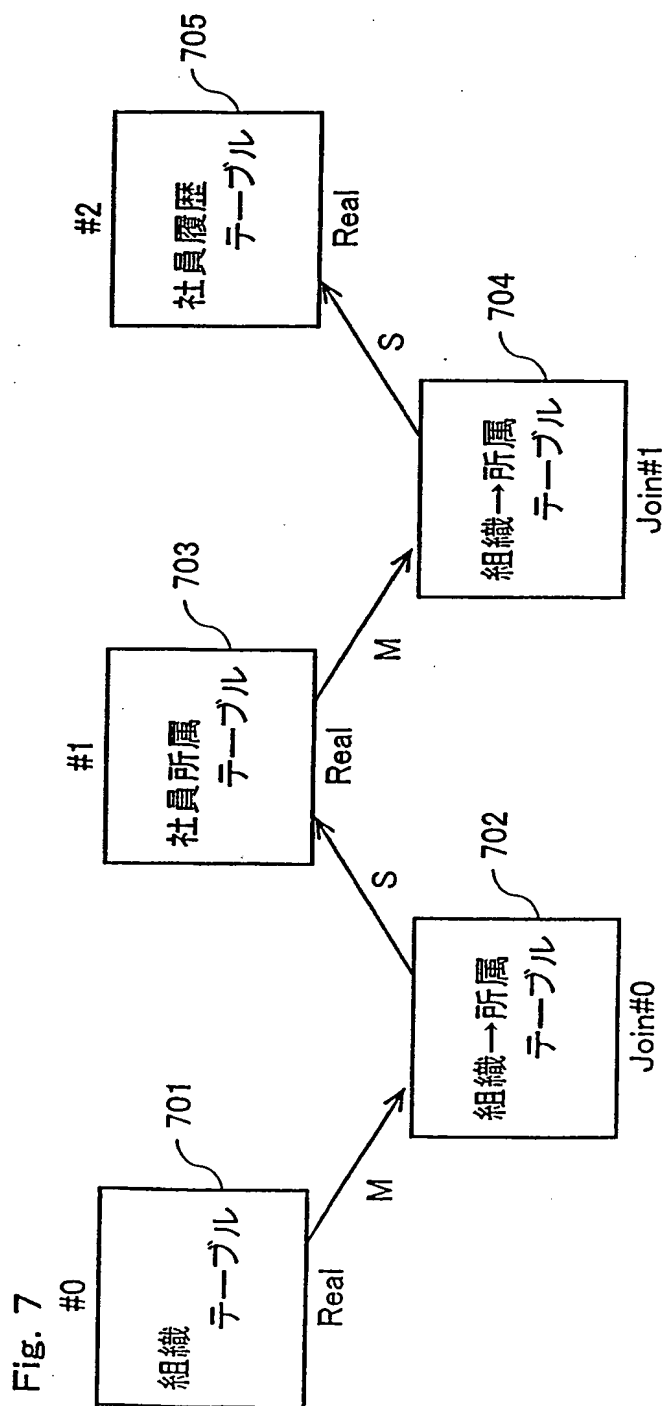


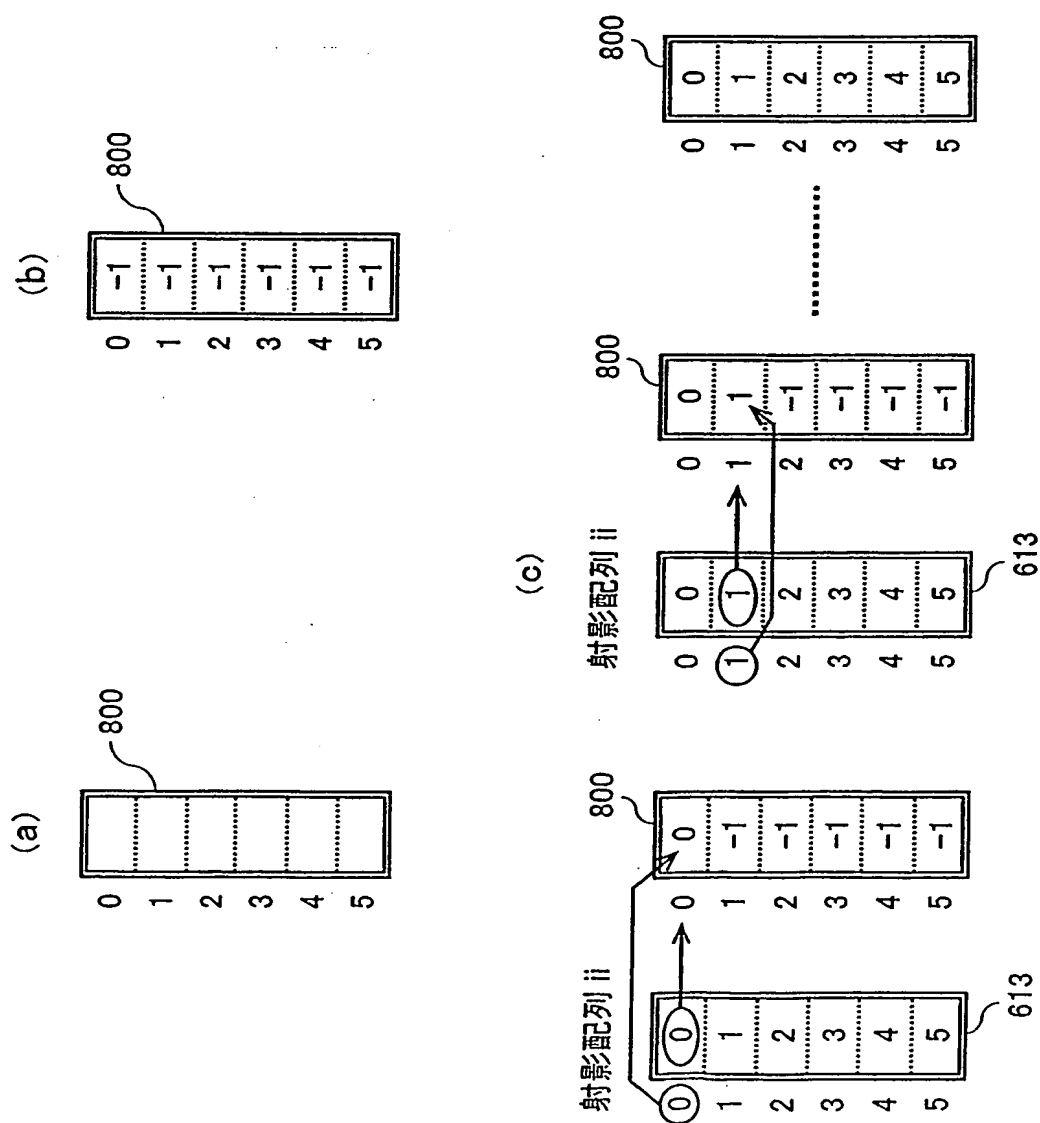
Fig. 5







உயிர்



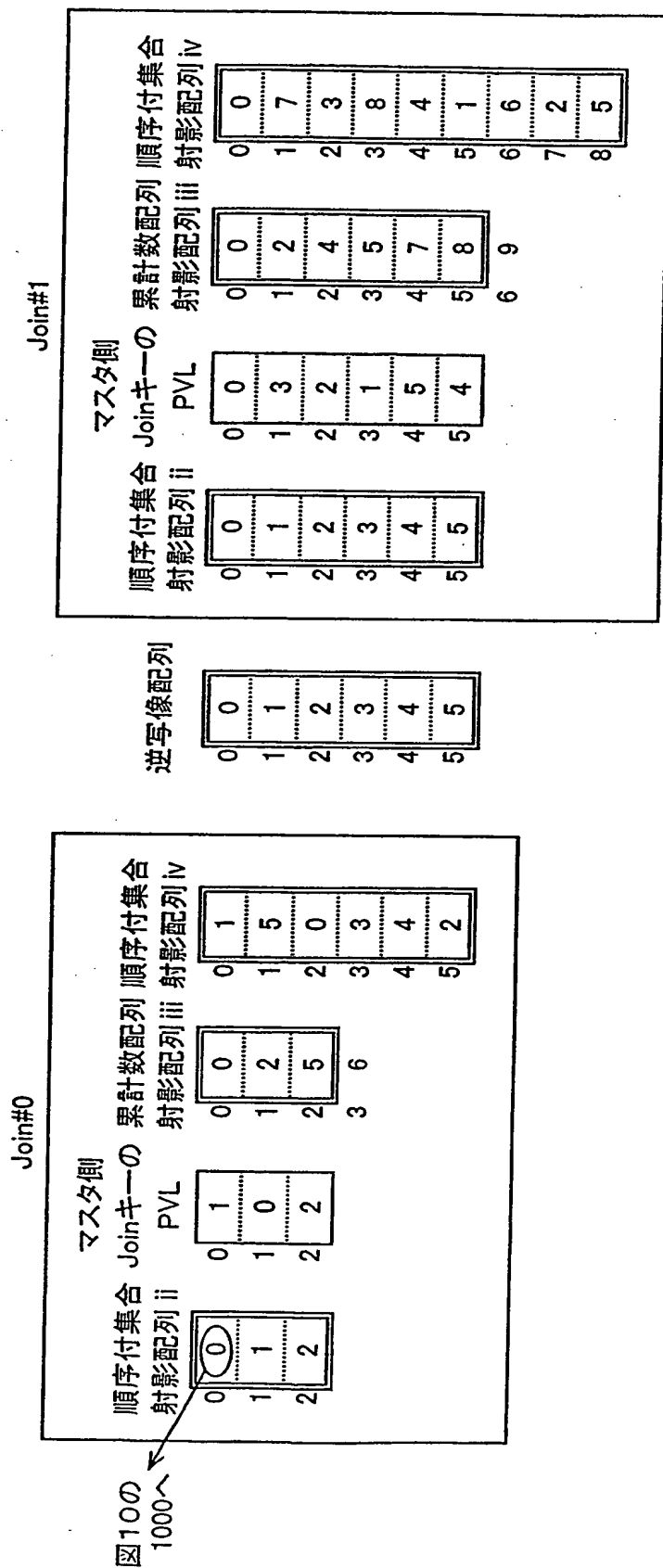
10/34

Fig. 10

900
✓

Depth	Rec No	
0	0	← 1000
1	0	← 1001
2	0	← 1002
2	7	← 1003
1	3	← 1004
2	3	← 1005
2	8	← 1006
1	4	← 1007
2	5	← 1008
0	1	← 1009
1	1	← 1010
2	1	← 1011
2	6	← 1012
1	5	← 1013
2	2	← 1014
0	2	← 1015
1	2	← 1016
2	4	← 1017

Fig. 11



1 2 / 3 4

Fig. 12

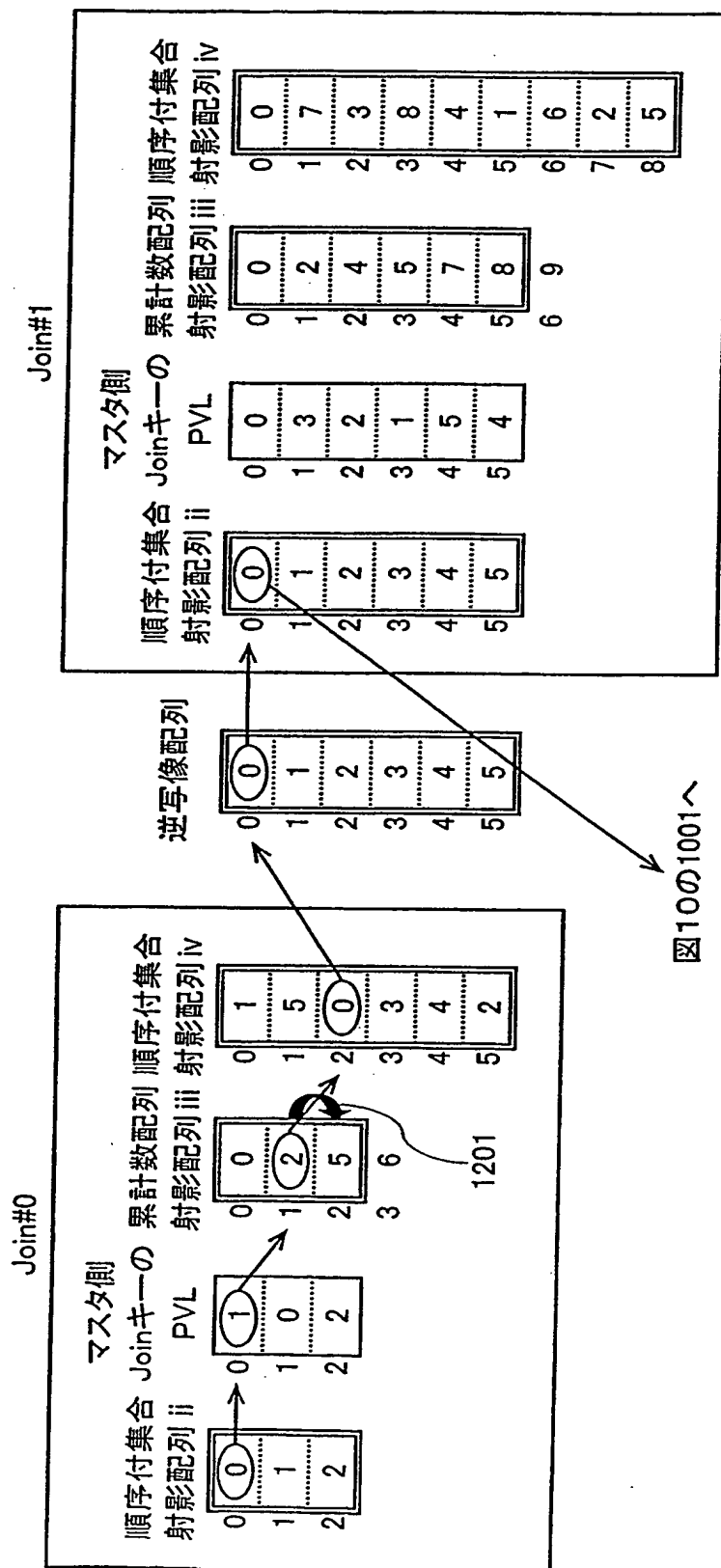
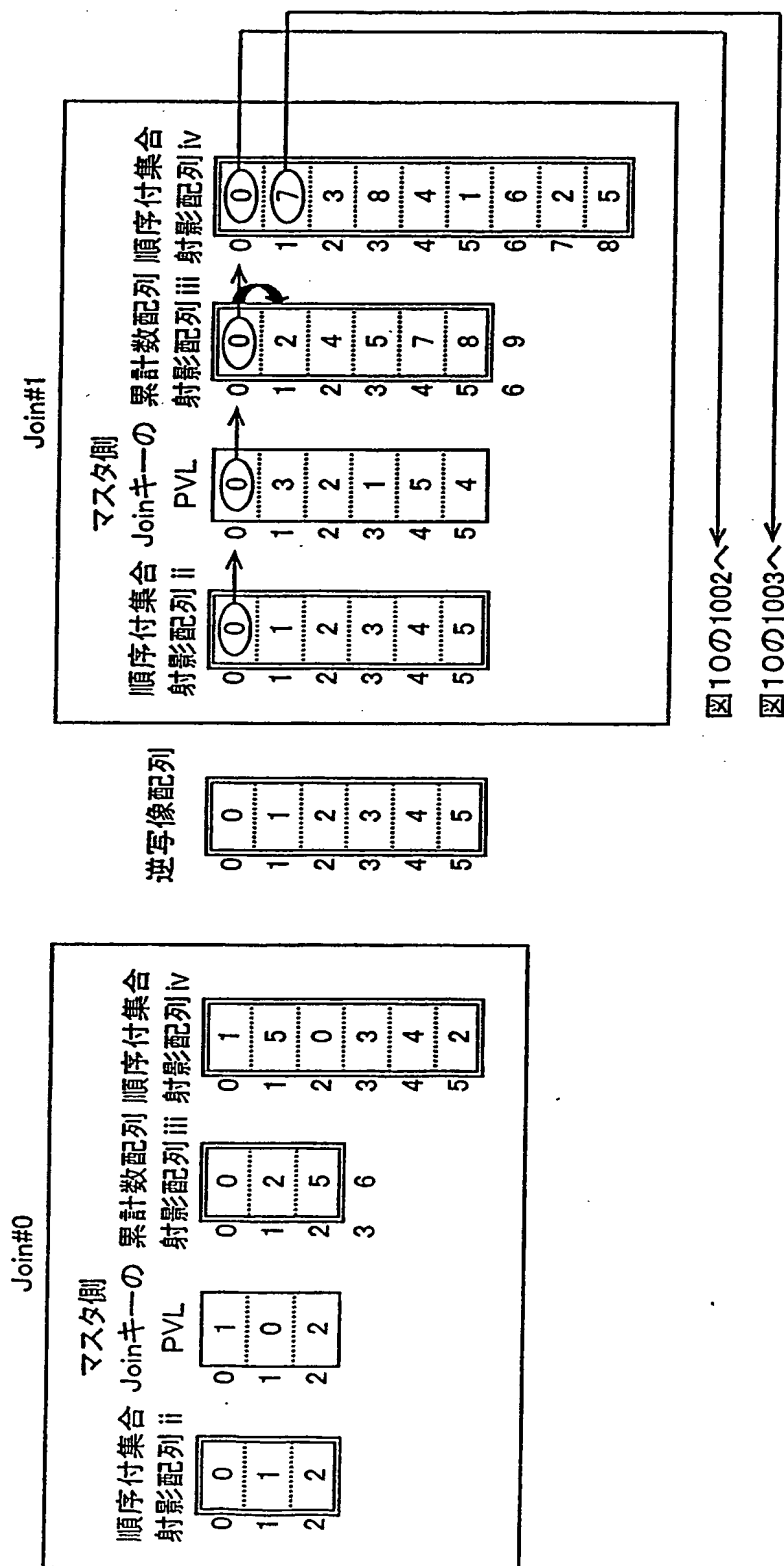


Fig. 13



14 / 34

Fig. 14

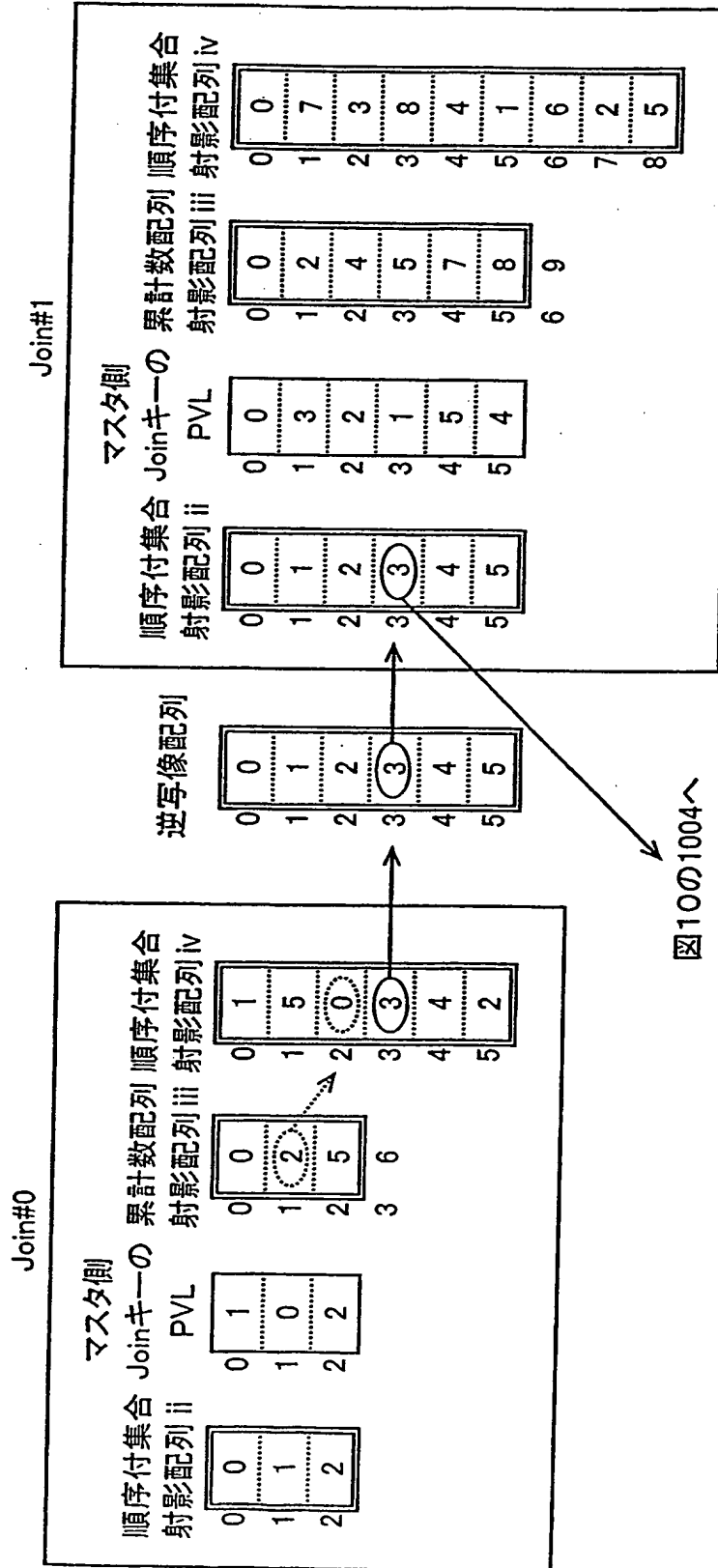


Fig. 15

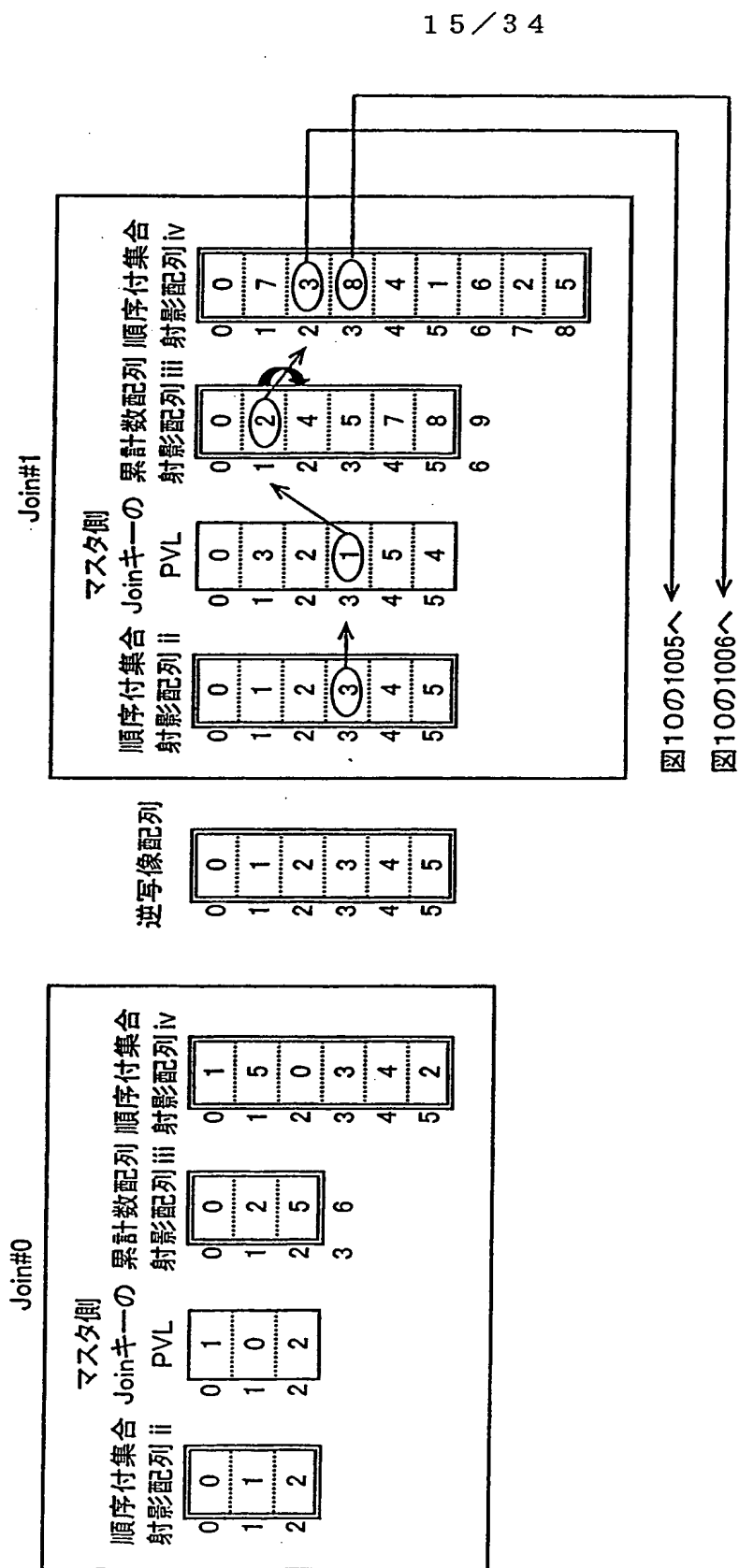


Fig. 16

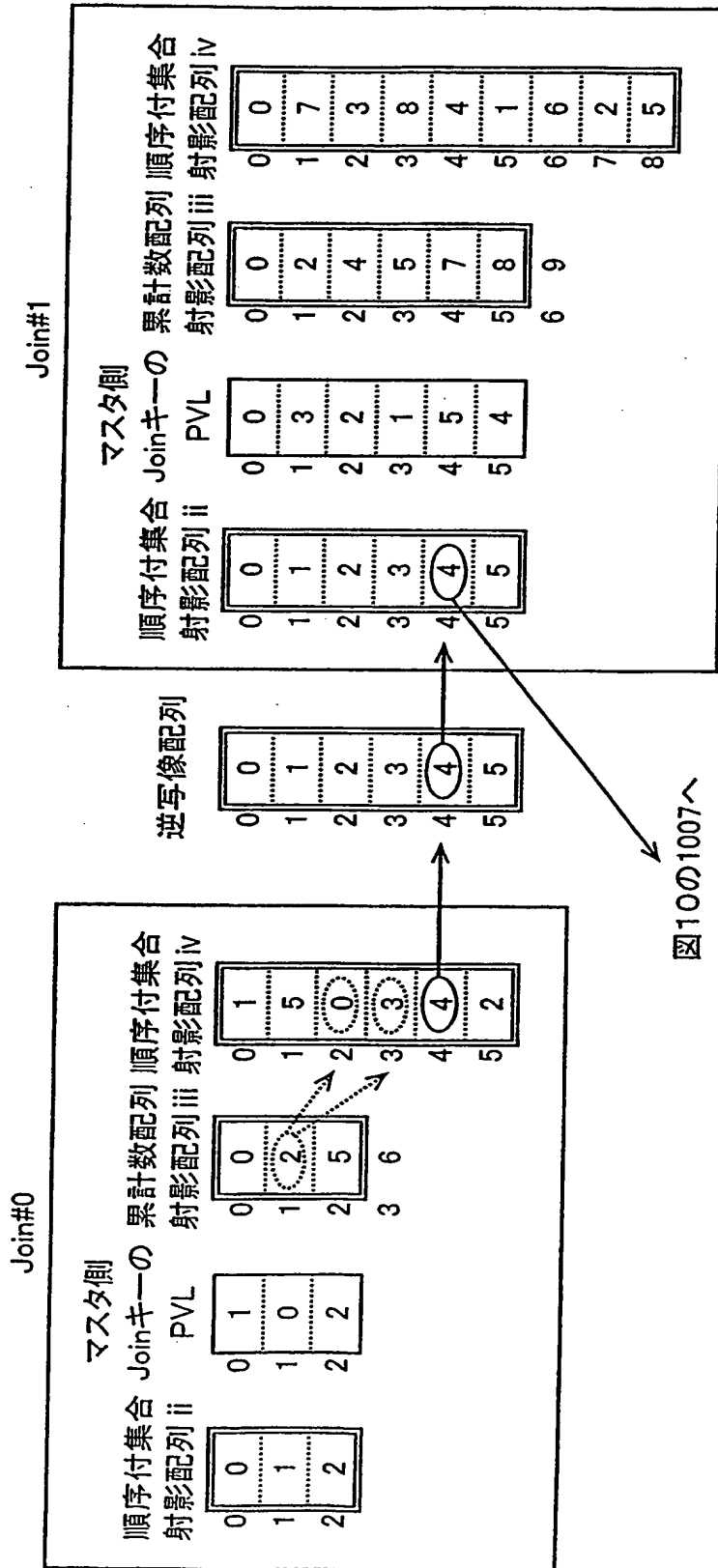


Fig. 17

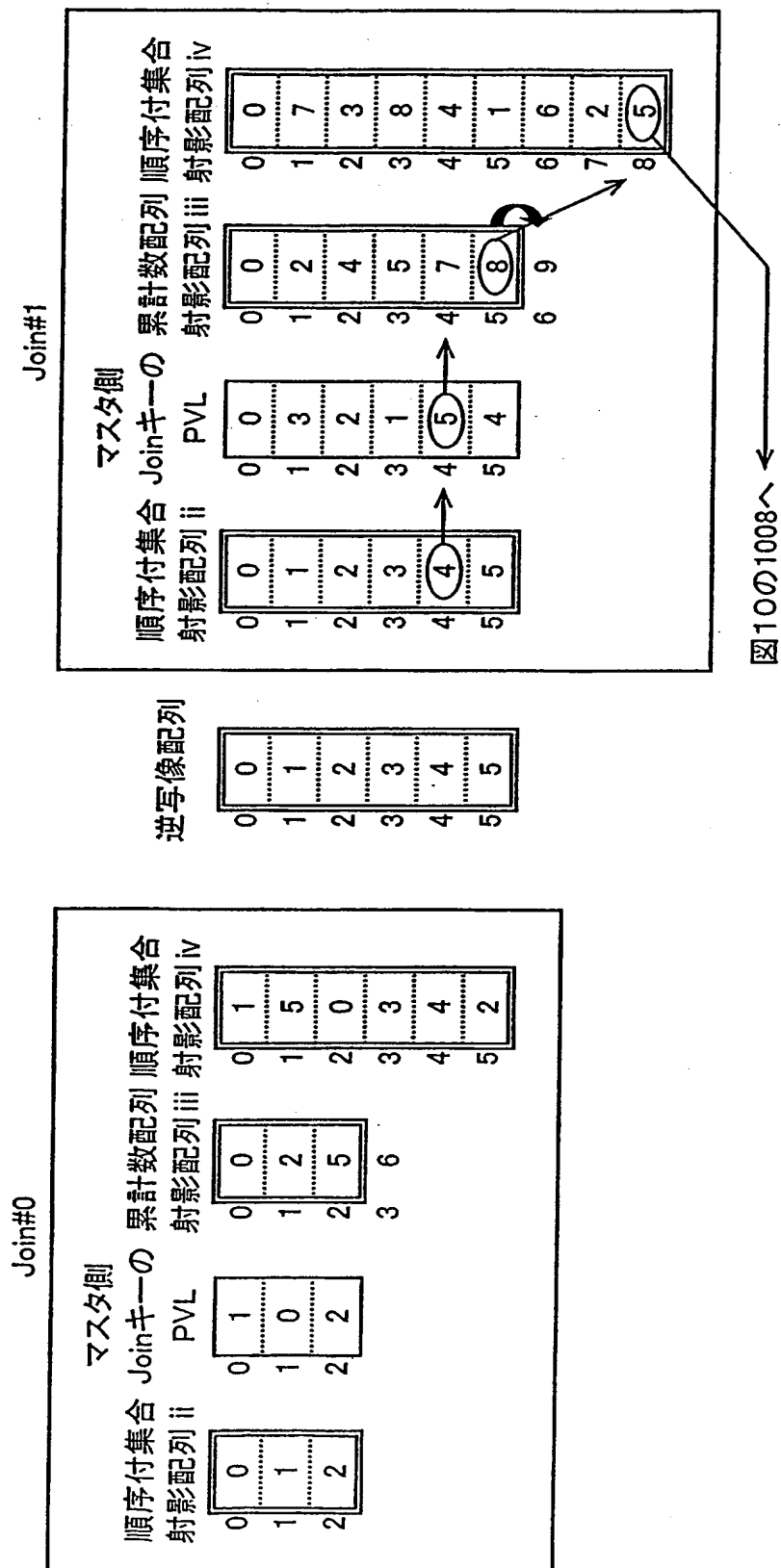


Fig. 19

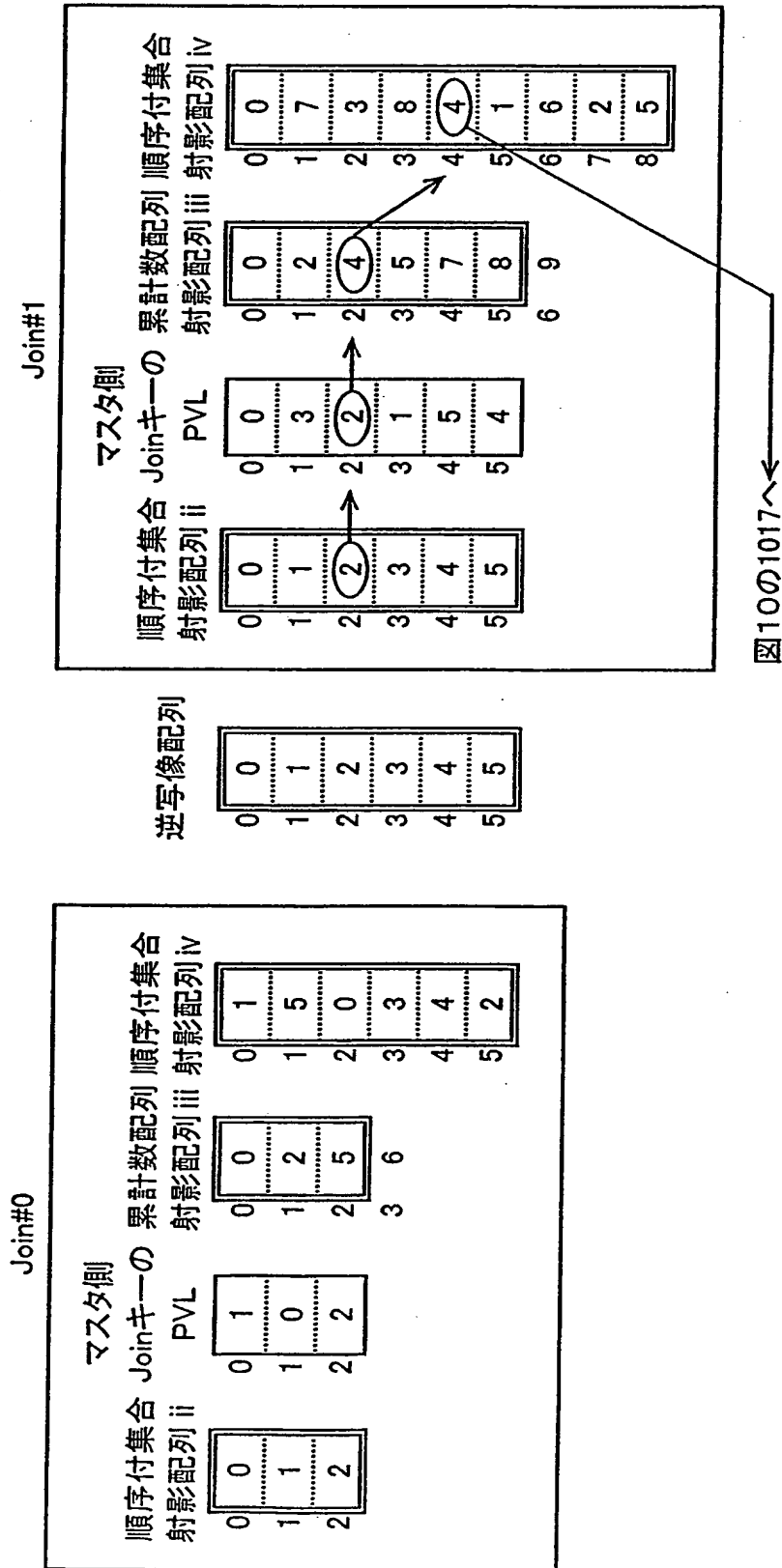


Fig. 20

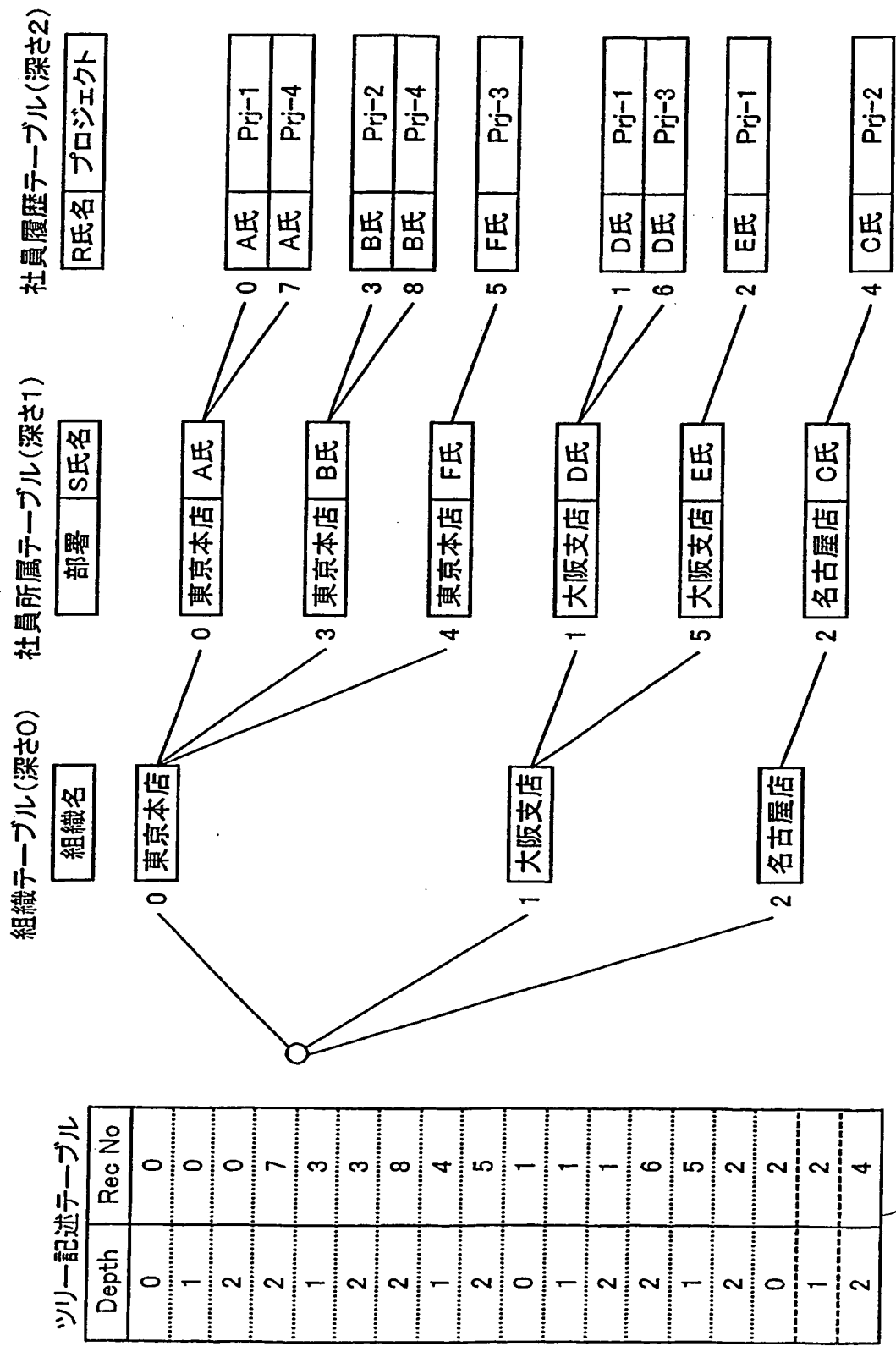


Fig. 21

組織テーブル

組織名
0 東京本店
1 大阪支店
2 名古屋店

(テーブル#0)

ソートなしの
全体集合

0 1
1 0
2 2

2101

組織名	
PVL	VL
0 1	0 大阪支店
1 0	1 東京本店
2 2	2 名古屋店

201

Fig. 22

社員所属テーブル

部署	S氏名
0 東京本店	A氏
1 大阪支店	D氏
2 名古屋店	C氏
3 東京本店	B氏
4 東京本店	F氏
5 大阪支店	E氏

(テーブル#1)

ソートなしの
全体集合

0	0
1	1
2	2
3	3
4	4
5	5

2201

「大阪、名古屋」
で検索

0	1
1	2
2	5

2202

「S氏名で」
ソート

0	2
1	1
2	5

2203

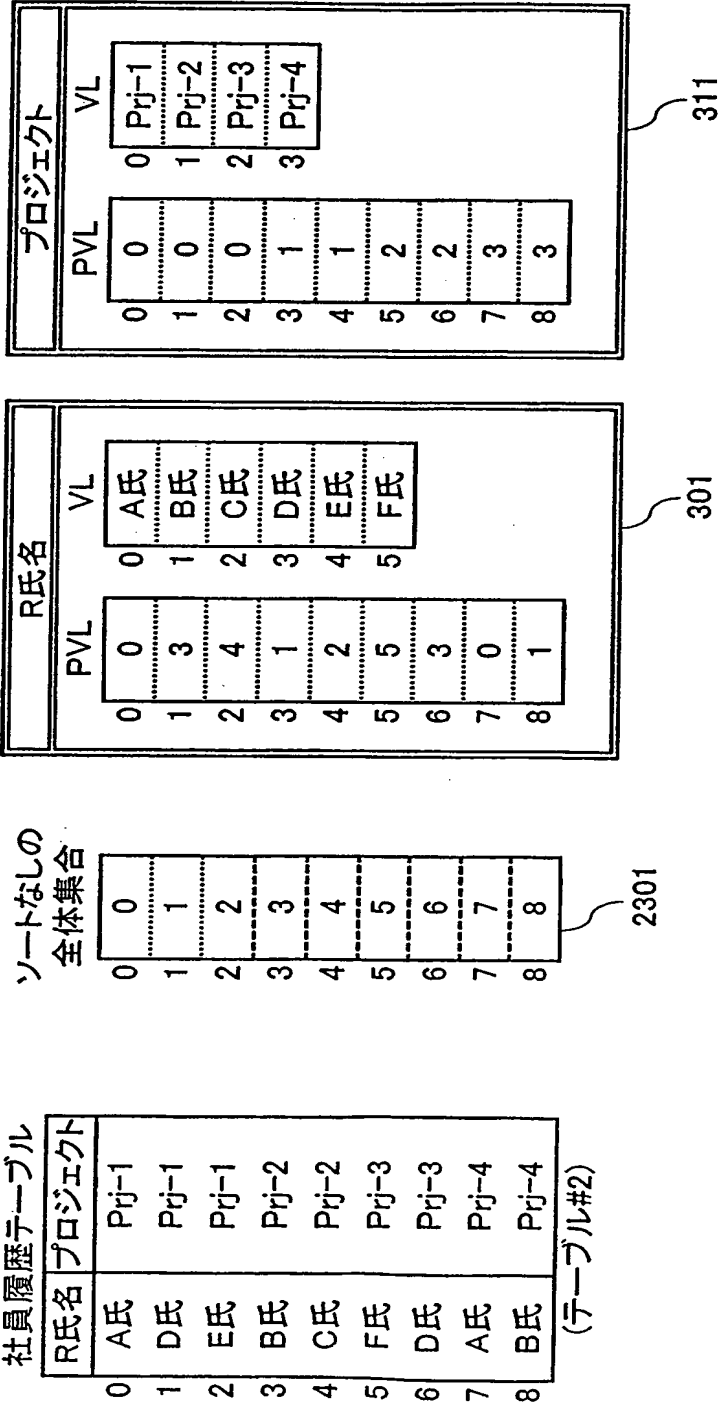
部署		PVL	VL
0	1	1	大阪支店
1	0	0	東京本店
2	2	2	名古屋店
3	1	1	
4	1	1	
5	0	0	

211

S氏名		PVL	VL
0	A氏	0	
1	B氏	3	
2	C氏	2	
3	D氏	1	
4	E氏	5	
5	F氏	4	

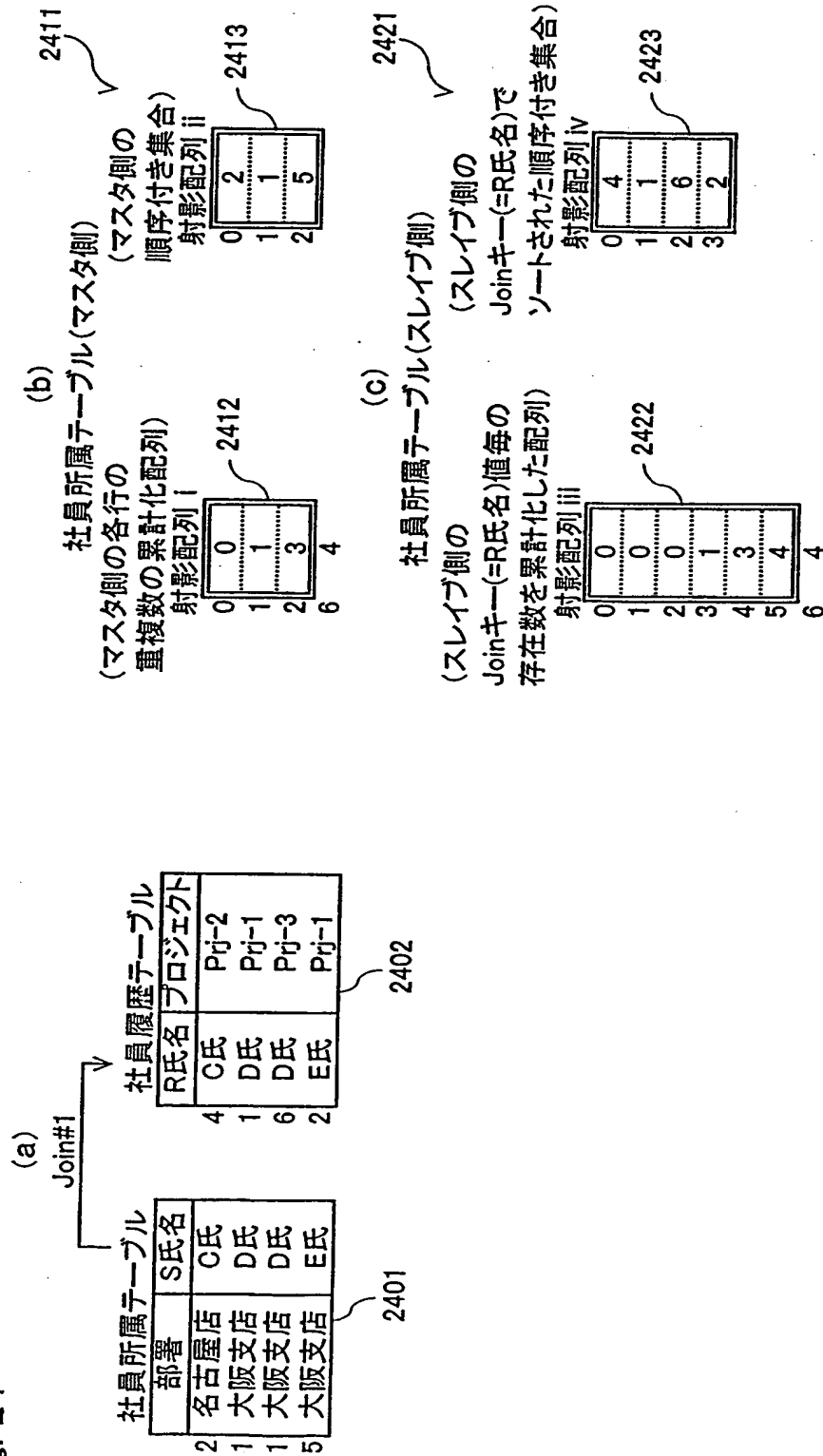
221

Fig. 23



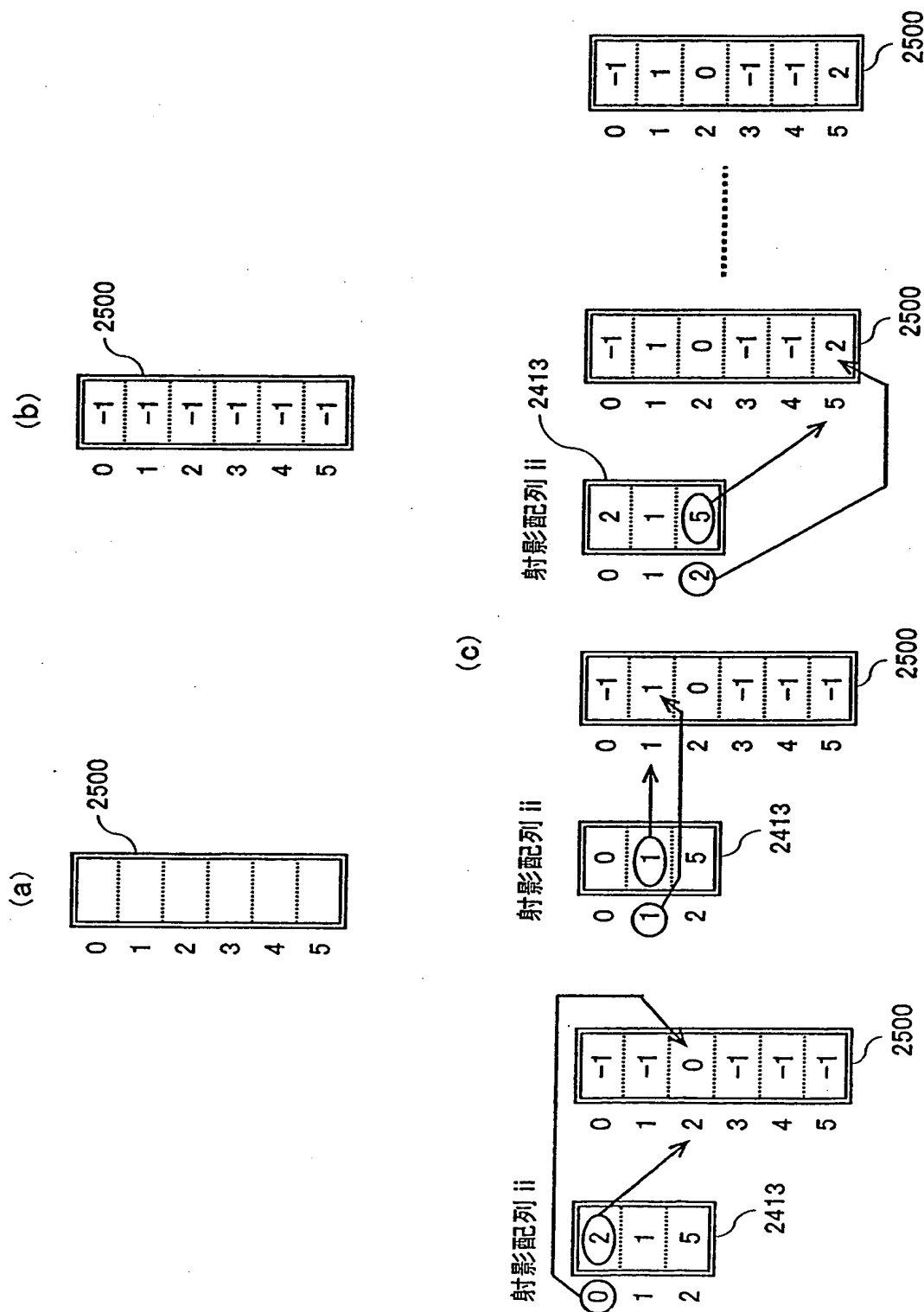
24 / 34

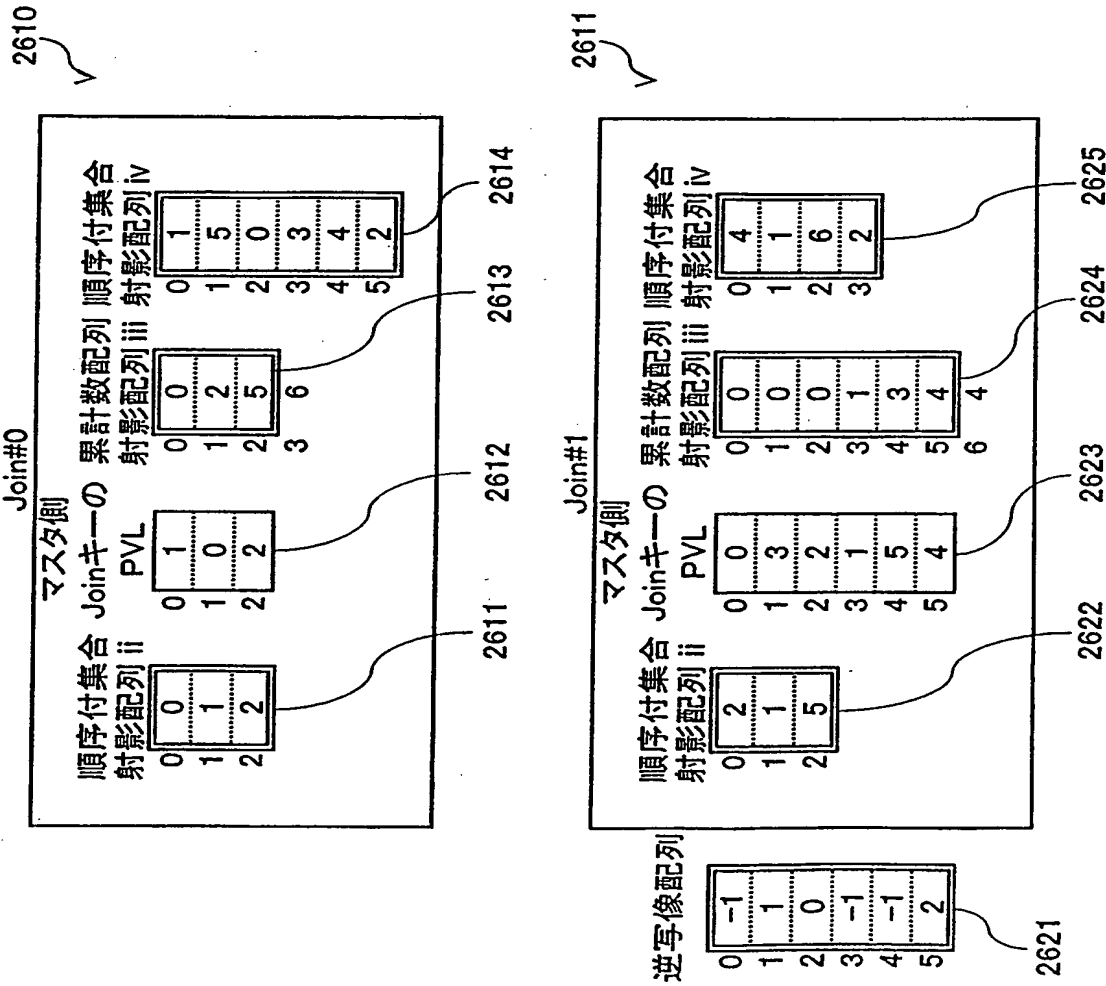
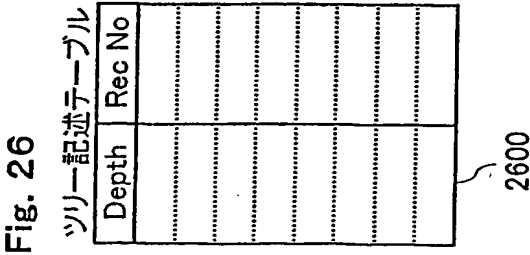
Fig. 24



25 / 34

Fig. 25





27 / 34

Fig. 27

2600
✓

Depth	Rec No	
0	0	← 2700
0	1	← 2701
1	1	← 2702
2	1	← 2703
2	6	← 2704
1	5	← 2705
2	2	← 2706
0	2	← 2707
1	2	← 2708
2	4	← 2709

Fig. 28

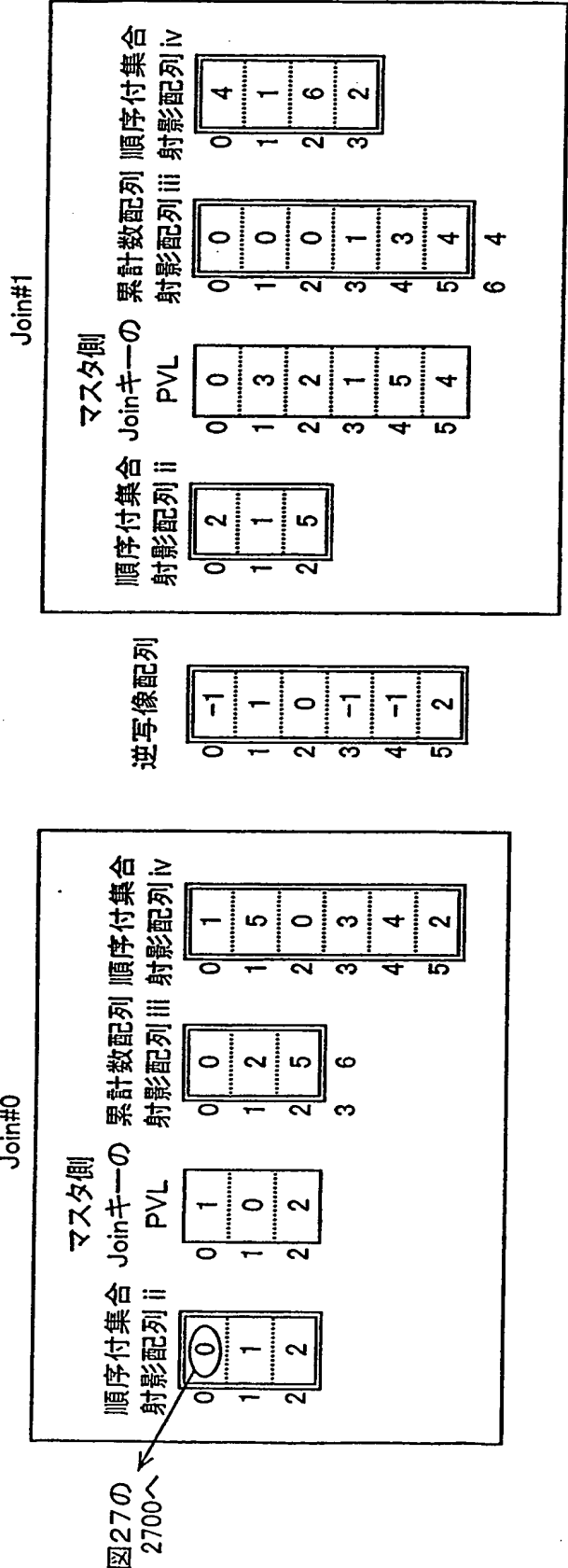
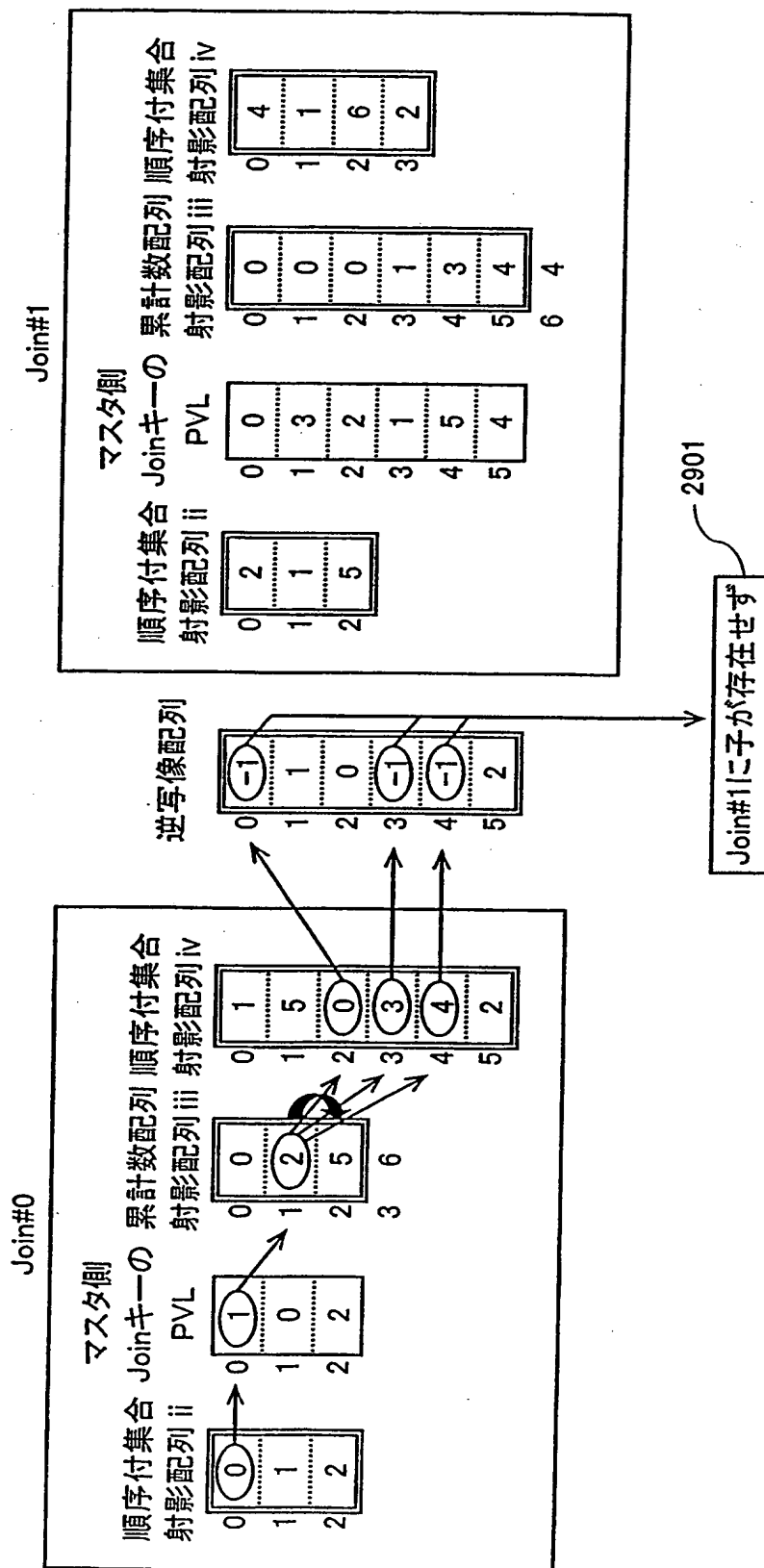


Fig. 29



30 / 34

Fig. 30

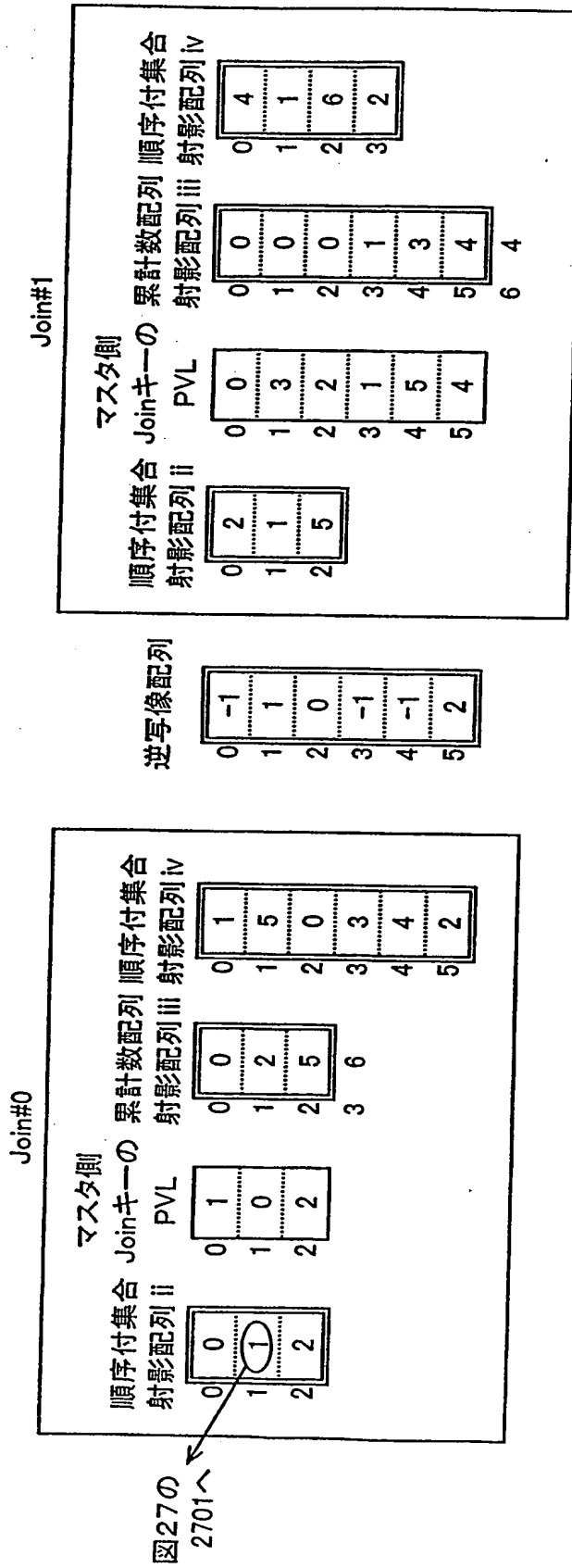
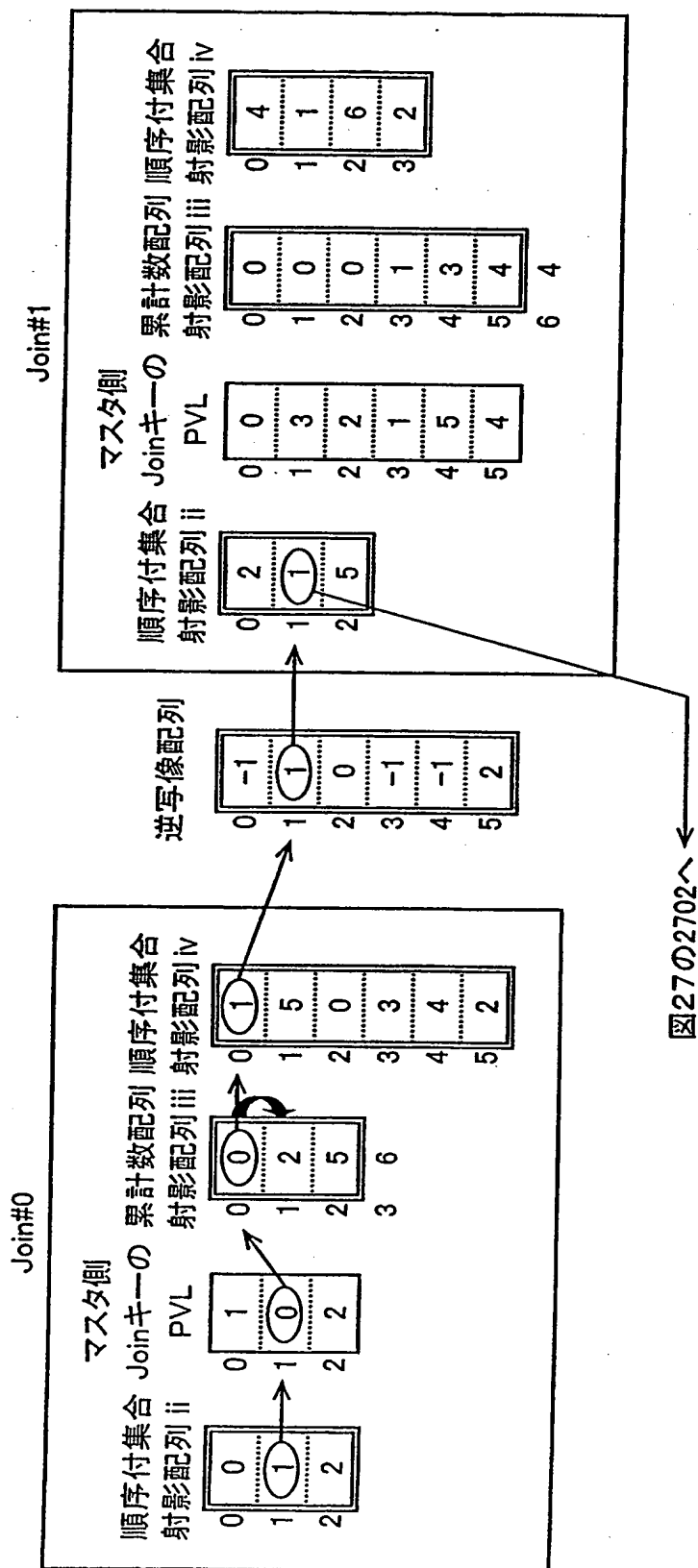


Fig. 31



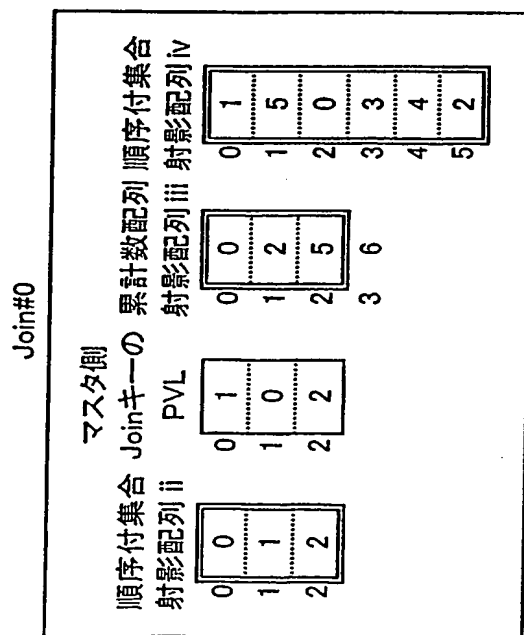
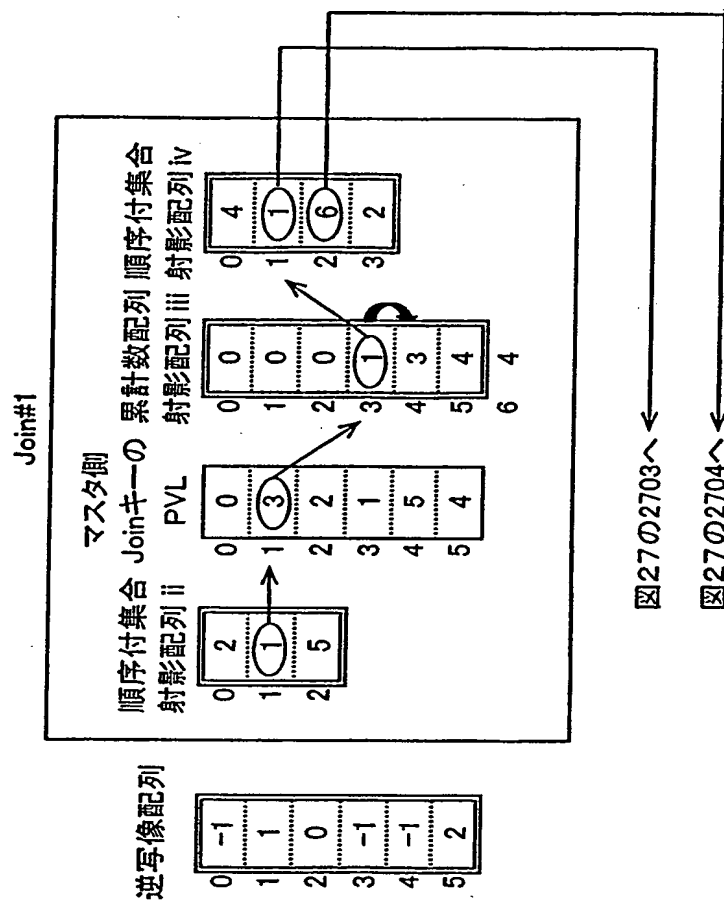
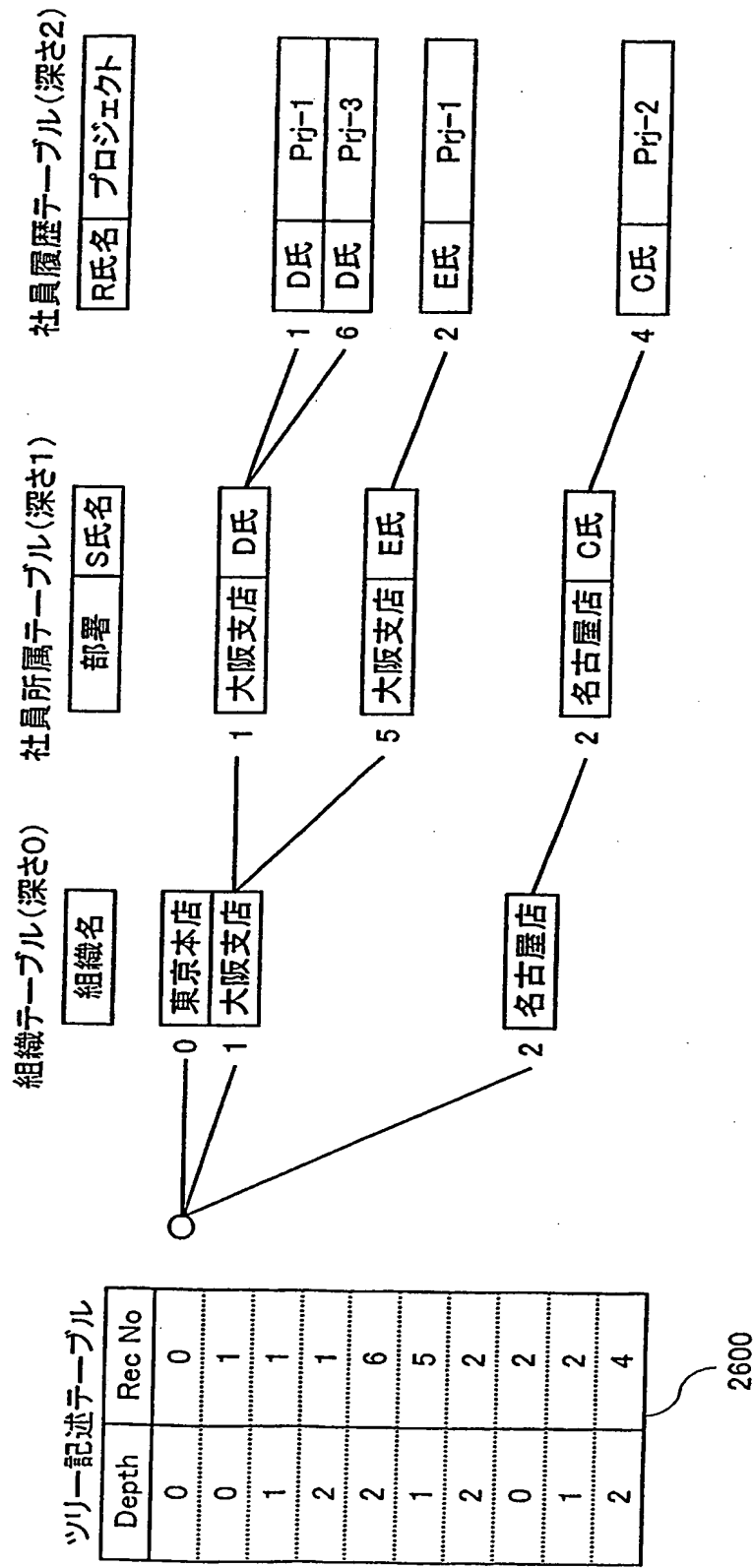


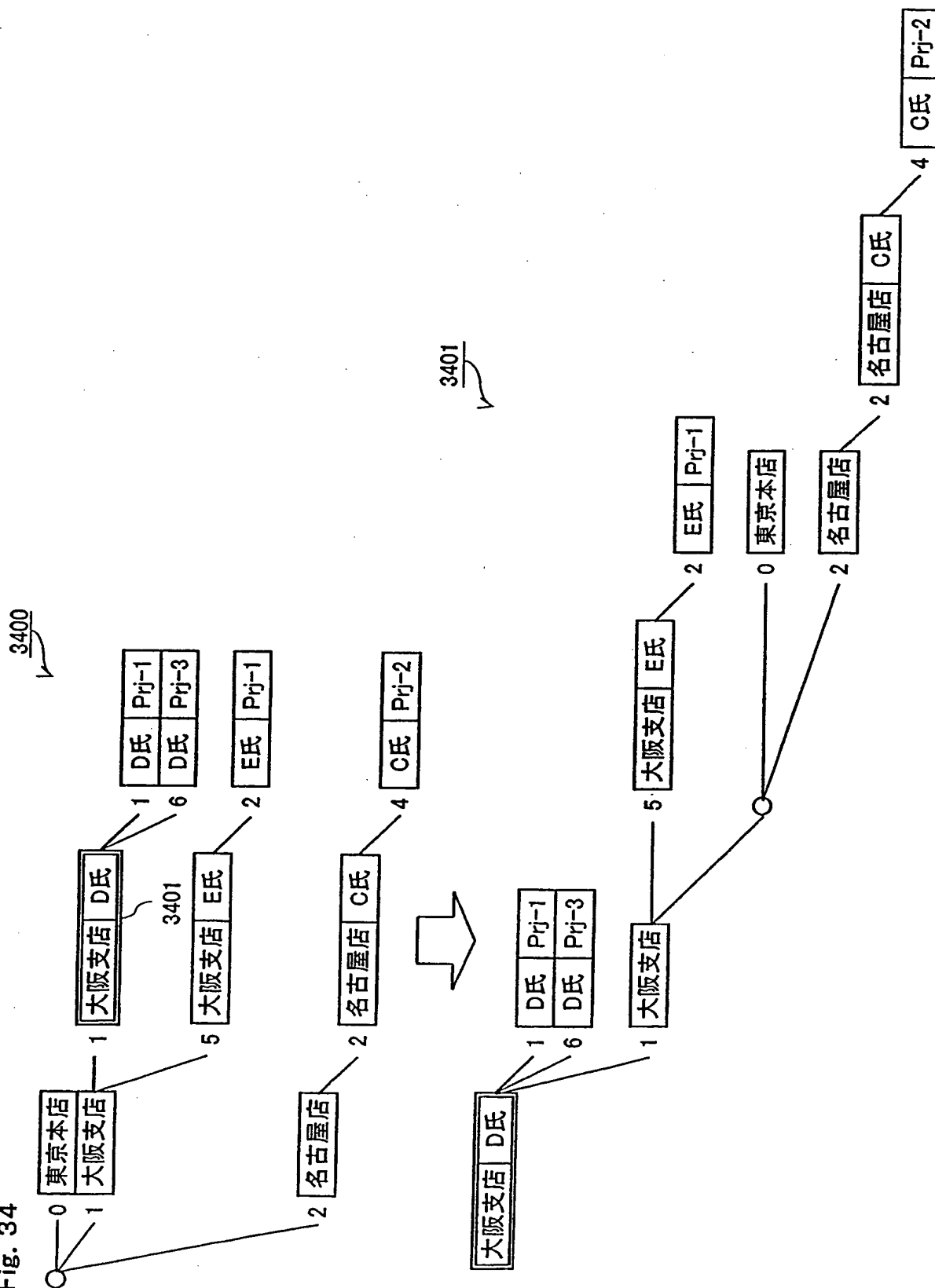
Fig. 32

Fig. 33



34 / 34

Fig. 34



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP03/13558

A. CLASSIFICATION OF SUBJECT MATTER
Int.Cl⁷ G06F17/30, G06F12/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl⁷ G06F17/30, G06F12/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Jitsuyo Shinan Koho 1922-1996 Toroku Jitsuyo Shinan Koho 1994-2004
Kokai Jitsuyo Shinan Koho 1971-2004 Jitsuyo Shinan Toroku Koho 1996-2004

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
JICST FILE (JOIS), WPI, INSPEC (DIALOG)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2000-339390 A (TURBO DATA LABORATORY INC.), 08 December, 2000 (08.12.00), & WO 00/73939 A & EP 1191462 A1 & CA 2375139 A & CN 1360700 T	1-10
A	WO 00/10103 A (Shinji FURUSHO), 24 February, 2000 (24.02.00), & EP 1136918 A1 & CA 2340008 A & CN 1317117 T	1-10
A	Norio DOI, "LFM o Tsukatta Kosoku In Memory XML Database Karearea, Database, No.21, 2002, April, pages 24 to 27	1-10

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search
26 January, 2004 (26.01.04)

Date of mailing of the international search report
10 February, 2004 (10.02.04)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

国際調査報告

国際出願番号 PCT/JP03/13558

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl⁷ G06F17/30, G06F12/00

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl⁷ G06F17/30, G06F12/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922-1996年
 日本国公開実用新案公報 1971-2004年
 日本国登録実用新案公報 1994-2004年
 日本国実用新案登録公報 1996-2004年

国際調査で使用了電子データベース (データベースの名称、調査に使用した用語)

JICSTファイル (JOIS), WPI, INSPEC (DIALOG)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	JP 2000-339390 A (ターボ・デ・ラボラトリー有限公司) 2000. 12. 08 &WO 00/73939 A &EP 1191462 A1 &CA 2375139 A &CN 1360700 T	1-10
A	WO 00/10103 A (古庄晋二) 2000. 02. 24 &EP 1136918 A1 &CA 2340008 A &CN 1317117 T	1-10
A	土井憲雄, LFMを使った高速インメモリXMLデータベースKarea rea, データベース, No. 21, 2002. 04, p. 24-27	1-10

☐ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの
 「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
 「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
 「O」 口頭による開示、使用、展示等に言及する文献
 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献
 「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
 「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
 「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
 「&」 同一パテントファミリー文献

国際調査を完了した日

26. 01. 2004

国際調査報告の発送日

10. 2. 2004

国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

高瀬 勤

5M

9069

電話番号 03-3581-1101 内線 3597